

Thompson, Chris

AI Task Force meeting - Aug 2, 2019 - 9:30 - 10:30

① Introductory remarks

- next steps: do something more concrete, e.g. publishing

s.21(1)(a)

s.21(1)(b)

② Follow-up on ACT project

- launch of CyberJustice program; we attended

- description of 16 research projects has arrived - Richard will forward

- TOPO comes in your sector for interest

- ~~prere~~ examples: paper/court support tools
pilot project with consumers
tools for self-represented litigants
court administration services on platform
inventory of best practices
ethical issues regarding AI

(how societies are looking at this)

IP and open data

(for the next mandate, there may be a revision of the Privacy Act)

- indicate if you would like to get involved in these groups

- we had committed to provide some lawyer time (in iCase) rather than money to support ACT

- will last for four years

- ③ Debrief on DM retreat

s.21(1)(a)

s.21(1)(b)

Page 4

**is withheld pursuant to sections
est retenue en vertu des articles**

21(1)(a), 21(1)(b)

**of the Access to Information Act
de la Loi sur l'accès à l'information**

s.21(1)(a)

s.21(1)(b)

- Q: Is there anyone from CSE invited?

They have some expertise in technology ... I wouldn't
doubt they have AI background too.

Also CSIS could be useful.

↓

Richard: We should look into this. This is in line
with the DM vision; we want to leverage
talents.

Richard: We could invite some of these people. One of the goals of the task force is to build information.

Dybold: One of the challenges with AI is protecting privacy.

Q/Q/Comment: These people piece together intelligence info.

① - Ethical use of algorithms.

② - Ethical decision-making.

s.21(1)(a)

s.21(1)(b)

Richard: Maybe we need a committee at a lower level. (Lower than DM...)

[Richard asked the NLS guy to study it for James to connect it with CSE.]

Q: Is there thought of looking at what we have?

A: Dybold:

but this is a good idea.

Richard: There is some EC discussion on tech resources.

Q: Is there work on policy for others?

↳ Federation of Law societies is having an AI conference in the fall.

↳ What is the government doing?

A: NLS: TBS is in the process of developing a Directive on Automated Decision-Making. (Spring 2019 gov)

Comment: last time I tried straight we couldn't
put comments into it

Dy. M: ~~We were~~ We're working on this.

Richard: We'll have a lot of training for Lex as well.

Dy. M: The goal is usability.

ent. F. H. H.
Noble:
Julie:

At the last meeting we presented the terms of reference

→ We'll reiterate the proposed terms of reference and discuss
the comments.

Richard: The working group ~~is~~ is a task force to be an
issue-focused.

Noble: It also includes a list of possible ^{members} ~~requirements~~.

↳ Noble can respond.

Richard: For the next meeting, we'll adopt

↳ Next meeting will be end of Aug. or beginning of
Sept?

↳ We'll look into more regular meetings.

s.21(1)(a)

s.21(1)(b)

s.23

- They will have to fill out a questionnaire.

- [REDACTED]

- Q: I can see people in the department coming to us in this.

Q: David Suskind in the UK.

A: The guy
- We've gathered tons of research - there's a lot of work being done at Harvard. It evolves quickly.

↳ e.g. Black box problem - lawyers have trouble explaining

↳ In the standard, you'll see certain criteria

↓
e.g. is something discrediting or misleading, etc.

↳ They're consulting inside and outside government on this tool.

↑
- Les Walden

F.H.H.:
- The 180 page report of the House of Lords is a good reference.

Julia:

[REDACTED]

Michelle Mann also wrote a summary of issues.

Richard: We need to have a shorthand site so all this is accessible.

Q: How do we disseminate information?

A: It's a problem we're looking at. We bought a tool called WalkMe to demonstrate.

- (5) Employment Forefight

- ↳ being implemented by Blue J for federal law
- ↳ we're supporting the development of a similar tool with PSPC

- (6) RFI support

- (7) Renable

- ↳ cloud-based update

- ↳ - there is a Community Cloud being set up
 - working through shared services is challenging
 - we're anticipating movement to the public cloud (1.8 million people for many to it for Office)

- we're now looking at cloud brokers for workloads (Symantec charges per person and per application)

- ↳ TBS also launched an RFI (Not Content
 - i) product solutions and services
 - ii) services
 - iii) solutions - TBS is involved in this)

- ↳ Dugald provided some ~~updates~~ comments regarding what Justice could do.

- ↳ Dugald: The procurement method is really used. There is movement in PSPC, but it's unclear.

N.H.L.:

s.21(1)(a)

s.21(1)(b)

s.23

↓
the new Minister of Digital Government has resulted
in numerous requests to be engaged

↓
in the context of the service strategy, we put together
a network; recently we're seeing linkages between
different projects... open governments etc.

↓
I'll find a way to send updates on this.

BC ← EMD

Page 11

**is withheld pursuant to section
est retenue en vertu de l'article**

23

**of the Access to Information Act
de la Loi sur l'accès à l'information**

s.19(1)

s.21(1)(a)

Thompson, Chris

s.21(1)(b)

From: Fobes.Caroline <Caroline.Fobes@cic.gc.ca> **s.23**
Sent: 2018-Aug-02 9:18 AM
To: Thompson, Chris; Tardif, Richard
Cc: Lafleur, Linda; IRCC.F SGCO Legal Services Unit / Services Juridiques BAGP F.IRCC
Subject: RE: AI/ML RFI update at AI Taskforce Meeting on August 2
Attachments: [REDACTED]

Chris and Richard, I am leaving at 2 pm [REDACTED] and all my meetings have been pushed to this morning so I will not be able to attend – my apologies.

[REDACTED]

Page 16

**is withheld pursuant to sections
est retenue en vertu des articles**

21(1)(a), 21(1)(b), 23

**of the Access to Information Act
de la Loi sur l'accès à l'information**

Caroline Fobes

s.21(1)(a)

s.21(1)(b)

Executive Director & Senior General Counsel, Legal Services (IRCC)
Immigration, Refugees and Citizenship Canada / Government of Canada
Caroline.Fobes@cic.gc.ca / Tel: 613-437-6722

Directrice executive & Avocate générale principale, Services juridiques (IRCC)
Immigration, Réfugiés et Citoyenneté Canada / Gouvernement du Canada
Caroline.Fobes@cic.gc.ca / Tél.: 613-437-6722

From: Thompson, Chris [mailto:Chris.Thompson@justice.gc.ca]

Sent: July 31, 2018 5:55 PM

To: Fobes.Caroline <Caroline.Fobes@cic.gc.ca>

Subject: AI/ML RFI update at AI Taskforce Meeting on August 2

Hi Caroline,

Regards,
--Chris

Thompson, Chris

From: Thompson, Chris
Sent: 2018-Aug-02 9:29 AM
To: Fobes.Caroline; Tardif, Richard
Cc: Lafleur, Linda; IRCC.F SGC Legal Services Unit / Services Juridiques BAGP F.IRCC
Subject: RE: AI/ML RFI update at AI Taskforce Meeting on August 2

s.21(1)(a)

Hi Caroline,

s.21(1)(b)

Thanks [REDACTED]

s.23

--Chris

From: Fobes.Caroline [mailto:Caroline.Fobes@cic.gc.ca]
Sent: 2018-Aug-02 9:18 AM
To: Thompson, Chris <Chris.Thompson@justice.gc.ca>; Tardif, Richard <Richard.Tardif@justice.gc.ca>
Cc: Lafleur, Linda <Linda.Lafleur@justice.gc.ca>; IRCC.F SGC Legal Services Unit / Services Juridiques BAGP F.IRCC <IRCC.SGCOLegalServicesUnit-ServicesJuridiquesBAGP.IRCC@cic.gc.ca>
Subject: RE: AI/ML RFI update at AI Taskforce Meeting on August 2

s.19(1)

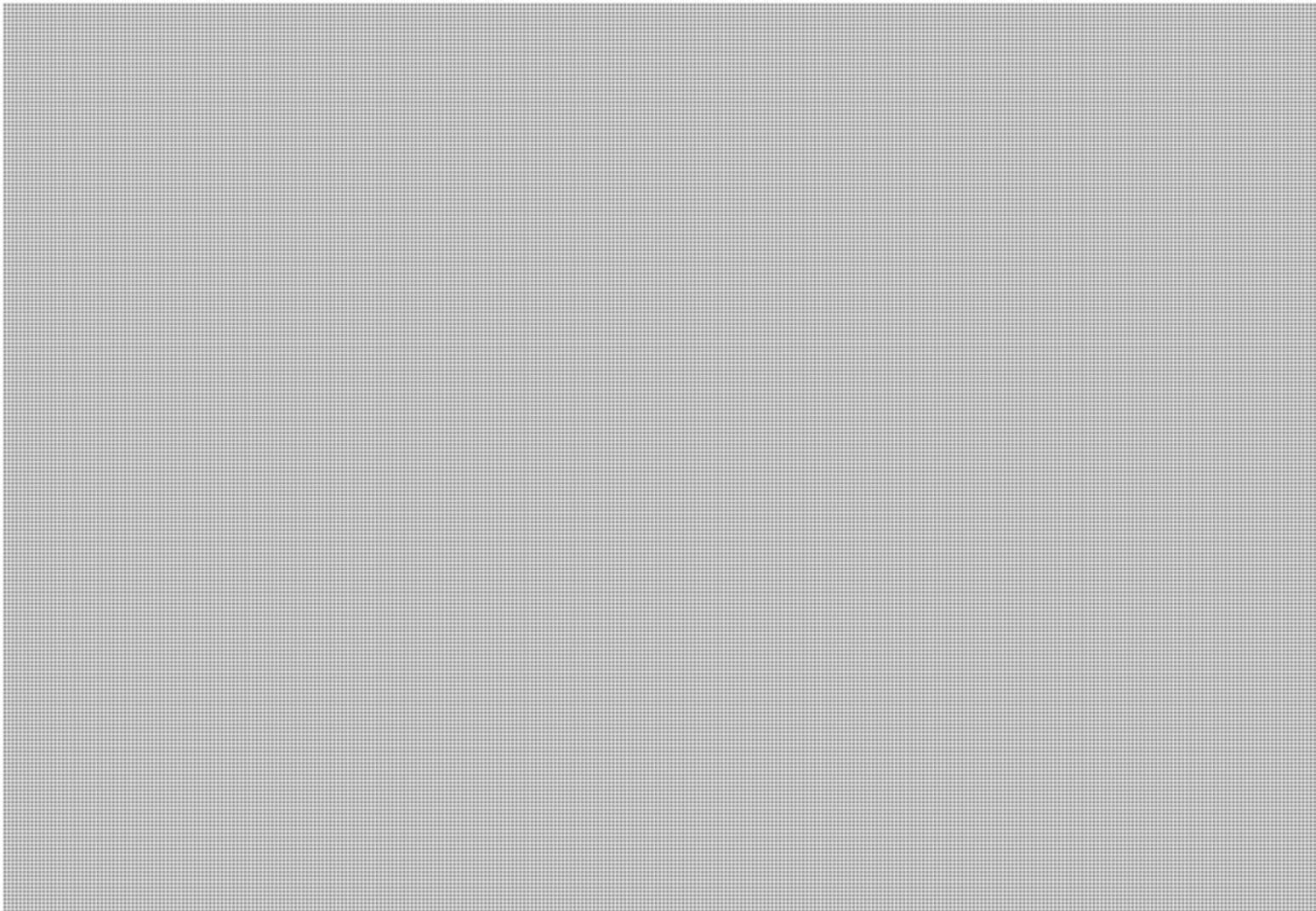
Chris and Richard, I am leaving at 2 pm [REDACTED] and all my meetings have been pushed to this morning so I will not be able to attend – my apologies.

Page 19

**is withheld pursuant to sections
est retenue en vertu des articles**

21(1)(a), 21(1)(b), 23

**of the Access to Information Act
de la Loi sur l'accès à l'information**



Caroline Fobes

s.21(1)(a)

s.21(1)(b)

Executive Director & Senior General Counsel, Legal Services (IRCC)
Immigration, Refugees and Citizenship Canada / Government of Canada
Caroline.Fobes@cic.gc.ca / Tel: 613-437-6722

s.23

Directrice exécutive & Avocate générale principale, Services juridiques (IRCC)
Immigration, Réfugiés et Citoyenneté Canada / Gouvernement du Canada
Caroline.Fobes@cic.gc.ca / Tél.: 613-437-6722

From: Thompson, Chris [<mailto:Chris.Thompson@justice.gc.ca>]

Sent: July 31, 2018 5:55 PM

To: Fobes.Caroline <Caroline.Fobes@cic.gc.ca>

Subject: AI/ML RFI update at AI Taskforce Meeting on August 2

Hi Caroline,



Regards,
--Chris

s.21(1)(a)

s.21(1)(b)

Thompson, Chris

From: Sarault, Nathalie <Nathalie.Sarault@tbs-sct.gc.ca>
Sent: s.21(1)(a) 2018-Aug-09 9:27 PM
To: s.21(1)(b) Tardif, Richard; McCombs, Mark (HRSDC); Zimmerman, Corinne; Wojcik, Nicolas;
s.23 Stringham, James; Wellington, Julie; Fobes, Caroline (CIC); Mann, Michelle P. (CIC);
Wong, Robert (FIN); McDonald, Susan; Reaney, Jennifer; Signorini, Gabriella; Zadro,
Matthew; Dubrule, Louis-Philippe; Thompson, Chris; Topshee, Dugald
Cc: Betts, Julia (TBS); Sarault, Nathalie (TBS)
Subject: RE: AI Working Group draft questions in relation to the AI WG Terms of Reference
Attachments: [REDACTED] Service Strategy -
Standing DOJ meetings; Service Strategy - Legal Project Management - August 9 2018 -
DOJ 8773000.docx; Annex A ENG White Paper.docx; Treasury Board Standard on
Automated Decision Making.docx; AI WG to the Task Force Terms of Reference August
2 2018 draft for Taskdocx

Hi everyone,

As promised at the last meeting, you will find enclosed:

- [REDACTED]
- The Legal Project Management Plan on the Service Strategy (evergreen).
- A recent e-mail sent to DOJ counsel involved in the Service Strategy – setting out linkages with other projects.
- The AI White Paper.

- [REDACTED]

p.s. just a friendly reminder to provide comments on the AI WG ToRs (re-attached for convenience, and see below e-mails). I will be on holidays Aug 13 to 31, so please ensure Michelle receives your comments as well.

Thanks,
Nathalie

Nathalie Sarault
613-907-7884

Senior Counsel, Treasury Board Secretariat Legal Services

Department of Justice Canada / Government of Canada
nathalie.sarault@tbs-sct.gc.ca / Tel: 613-907-7884 / TTY: 613-369-9371

Avocate-conseil, Service juridique du Secrétariat du Conseil du Trésor
Ministère de la Justice Canada / Gouvernement du Canada
nathalie.sarault@tbs-sct.gc.ca / Tél. : 613-907-7884 / ATS : 613-369-9371

From: Sarault, Nathalie

Sent: August 2, 2018 11:38 AM

To: 'Tardif, Richard' <Richard.Tardif@justice.gc.ca>; McCombs, Mark: HRSDC.RHDCC <mark.mccombs@hrsdc-rhdcc.gc.ca>; Zimmerman, Corinne: JUS.JUS <corinne.zimmerman@justice.gc.ca>; Wojcik, Nicolas: JUS.JUS <nicolas.wojcik@justice.gc.ca>; Stringham, James <James.Stringham@justice.gc.ca>; Wellington, Julie: JUS.JUS <julie.wellington@justice.gc.ca>; Fobes, Caroline (CIC) <Caroline.Fobes@cic.gc.ca>; Mann, Michelle: CIC.CIC <Michelle.Mann@cic.gc.ca>; Wong, Robert (FIN) <robert.wong@canada.ca>; McDonald, Susan: JUS.JUS <susan.mcdonald@justice.gc.ca>; Reaney, Jennifer <Jennifer.Reaney@justice.gc.ca>; Signorini, Gabriella: JUS.JUS <gabriella.signorini@justice.gc.ca>; Zadro, Matthew <Matthew.Zadro@justice.gc.ca>; Dubrule, Louis-Philippe: JUS.JUS <louis-philippe.dubrule@justice.gc.ca>; Thompson, Chris <Chris.Thompson@justice.gc.ca>; Topshee, Dugald <Dugald.Topshee@justice.gc.ca>

Cc: Sarault, Nathalie <Nathalie.Sarault@tbs-sct.gc.ca>

Subject: RE: AI Working Group draft questions in relation to the AI WG Terms of Reference

Hi everyone,

Following our meeting this morning, I am re-sending the below for your comments (for those who had not yet reviewed). It would be appreciated if you could provide comments by end of next week (copying myself, Michelle and Richard), but don't hesitate to let us know if that is not feasible.

I have also attached a slightly updated version of the ToRs taking into account comments provided to date.

Many thanks.

Sent on behalf of Nathalie and Michelle

Nathalie Sarault
613-907-7884

Senior Counsel, Treasury Board Secretariat Legal Services
Department of Justice Canada / Government of Canada
nathalie.sarault@tbs-sct.gc.ca / Tel: 613-907-7884 / TTY: 613-369-9371

Avocate-conseil, Service juridique du Secrétariat du Conseil du Trésor
Ministère de la Justice Canada / Gouvernement du Canada
nathalie.sarault@tbs-sct.gc.ca / Tél. : 613-907-7884 / ATS : 613-369-9371

From: Tanyan, Caroline [<mailto:Caroline.Tanyan@justice.gc.ca>] **On Behalf Of** Tardif, Richard
Sent: June 22, 2018 6:59 PM
To: McCombs, Mark: HRSDC.RHDCC <mark.mccombs@hrsdc-rhdcc.gc.ca>; Zimmerman, Corinne: JUS.JUS <corinne.zimmerman@justice.gc.ca>; Wojcik, Nicolas: JUS.JUS <nicolas.wojcik@justice.gc.ca>; Stringham, James <James.Stringham@justice.gc.ca>; Wellington, Julie: JUS.JUS <julie.wellington@justice.gc.ca>; Fobes, Caroline (CIC) <Caroline.Fobes@cic.gc.ca>; Mann, Michelle: CIC.CIC <Michelle.Mann@cic.gc.ca>; Wong, Robert (FIN) <robert.wong@canada.ca>; McDonald, Susan: JUS.JUS <susan.mcdonald@justice.gc.ca>; Reaney, Jennifer <Jennifer.Reaney@justice.gc.ca>; Signorini, Gabriella: JUS.JUS <gabriella.signorini@justice.gc.ca>; Zadro, Matthew <Matthew.Zadro@justice.gc.ca>; Dubrule, Louis-Philippe: JUS.JUS <louis-philippe.dubrule@justice.gc.ca>; Thompson, Chris <Chris.Thompson@justice.gc.ca>; Topshee, Dugald <Dugald.Topshee@justice.gc.ca>; Sarault, Nathalie <Nathalie.Sarault@tbs-sct.gc.ca>
Subject: AI Working Group draft questions in relation to the AI WG Terms of Reference

Dear AI Task Force members,

I would like to thank again Michelle Mann and Nathalie Sarault for preparing a draft of the Terms of Reference for the DOJ AI WG (latest draft is attached).

In order to finalize the TOR, they would benefit from our comments on the following points:

1. Whether **Mandate** should expressly be stated to include the following (new text to consider in green):
 - determining the need for and pursuing consultations with officials of other departments of the GoC , the private sector, academia, other international governments or organisations.

A specific example, should we be formally liaising with the CBA for example and suggesting that they set up a WG on AI? Mark McCombs provided very helpful views on these aspects where he noted that given the level of interest in this area and complexity associated with direct outreach to the private bar we should rather do so via a CBA WG and he was supportive of collaboration with academia – as I know are you. Based on this I have removed the reference below to outreach to law firms.

2. Further to this same Mandate question, whether **Responsibilities and Key Deliverables** should include reference to the following (new text to consider in green):
 - where appropriate, conduct research and, develop analyses including through collaboration with AI experts (public and private sector), innovation labs, governments of provinces or countries or organisations that are thought leaders in relation to AI technology to consider their perspectives on legal, legal policy and ethical considerations;
 - approach the Canadian Bar Association (CBA) to propose the formation of a CBA AI Working Group on AI to review legal and legal- policy considerations in relation to AI technologies.
3. Whether **Responsibilities and Key Deliverables** should include reference to the following or do members have other ideas in relation to the need to consult client departments:
 - Consult key client departments to ensure their input is considered in relation to deliverables developed such as frameworks, checklists etc.

We are suggesting that we add the above text in green to ensure that key client departments be consulted as work progresses and before decisions are made that may impact their work. While this has been done effectively on an ad hoc basis thus far by the Task Force in respect of the development of AI tools in the legal field, for the new AI WG and its deliverables this warrants some specific consideration as we move forward. To this end, we could develop a consultation strategy that is customised to the deliverable we are working on. For example, the need to consult client departments would vary

depending upon whether we are developing a legal framework tool vs a foundation opinion. However, members may have better ideas on how to address this point.

4. We are wondering what support/resources we might obtain to assist with this work in particular secretariat type functions (note taking, organising meetings, loading content to the DW and Study Group pages etc).
5. We took the liberty of proposing some members of the core WG based on practitioners who have started working in the area or tech law matters generally but of course this is entirely up to management.

I was wondering if you could provide Michelle and Nathalie with your comments on the TOR and the 5 questions prior to next Thursday. Thank you also for cc'ing me on your replies.

Regards.

Richard

M^e Richard L. Tardif

Director General and Senior General Counsel |
Directeur général et avocat général principal
Legal Practices Sector | Secteur des pratiques juridiques
Department of Justice Canada | Ministère de la Justice Canada
275 Sparks Street, 10th floor, SAT-10012 | 275, rue Sparks, 10^e étage, TSA-10012
Ottawa (Ontario) K1A 0H8
Tel. | Tél. : 613-952-3816
richard.tardif@justice.gc.ca

CONFIDENTIALITY NOTICE: The contents of this email are confidential and reserved for the sole use of the intended recipients. Any use, disclosure or copying of the information is strictly prohibited. This email may also contain information protected by solicitor-client privilege. If you receive this message in error, please notify the sender immediately and destroy the original message and all copies. Thank you.

AVIS: Ce courriel contient des renseignements confidentiels dont l'usage est réservé exclusivement à la personne à laquelle il est destiné. Toute utilisation, divulgation ou reproduction de cette information est strictement interdite. Ce courriel peut aussi contenir de l'information privilégiée qui est protégée par le secret professionnel de l'avocat-client. Si vous recevez cette communication par erreur, veuillez en aviser l'auteur immédiatement et détruire l'original et toute copie. Merci.

Thompson, Chris

From: Sarault, Nathalie <Nathalie.Sarault@tbs-sct.gc.ca>
Sent: 2018-Jun-22 5:00 PM
To: Clervoix, Magali; Horton, Andrea: JUS.JUS; Laframboise, Mélanie; Palmer, Janice (CRA); Deziel, Nathalie (HRSDC); Carruthers, Susan (VAC); Mann, Michelle P. (CIC); Van Rutten, Béatrice (IC); Lewis, Kimberley (PSC); Signorini, Gabriella
Cc: Betts, Julia (TBS); Bambrick, Lisa (TBS); Geh, Sarah; Caron, Daniel; Grace, Elizabeth (TBS); Steers, Isabelle (IC); Sarault, Nathalie (TBS)
Subject: Service Strategy - Standing DOJ meetings
Attachments: [REDACTED]

s.69(1)(g) re (e)

Hi everyone,

I'm writing to provide you with an update on the Service Strategy and give you a heads up about certain aspects.

- I will be [REDACTED] back on July 3; Julia will be here.
- As mentioned during the last call, I think our work on the Service Strategy will have close linkages with other projects that TBS and other departments are working on, namely:
 - OneGC – technological vision of the GoC operating as one; it includes DXP (digital exchange platform, formerly the interoperability platform) and “tell us once” – the concept of Canadians being able to provide information only once to the GoC rather than having to repeat it when dealing with different departments.
 - Digital Policy (includes enabling digital services to Canadians and the work on artificial intelligence).
 - Data Strategy - to better manager GC information as a valuable asset; it includes data relevant for providing services to Canadians, for statistical purposes, and for informing policy and legislative/regulatory decisions.

s.19(1)




s.21(1)(a)

s.21(1)(b)

s.23

• [REDACTED]

• [REDACTED]

- 
 TBS will then contact relevant departments and ask for further validation and input to inform the discussion on what challenges departments are facing. Our TBS clients have asked us if you know who are the key contacts in your departments regarding information-sharing agreements. They have their own contacts but they want to make sure they get to all the right people.
s.21(1)(a)
s.69(1)(g) re (e) s.21(1)(b)
- Key upcoming dates: s.23
 - 
 - July 13ish, 2018 – TBS to send departments a “Questions Paper” to inform next steps for the 3 “quick wins”, in preparation for a workshop
 - Early August – workshop with departments to discuss quick wins
 - August – tasking to departments to identify use cases (i.e. services requiring improvement)
 - September – December: in-depth review of use cases and plan for resolution (i.e. improved service state)
- Some of you have asked if there was an updated description of the Service Strategy. We’ll have to check with the clients for the most recent and comprehensive description, but meanwhile, here are some excerpts from older documents: s.69(1)(g) re (a)

Nathalie Sarault
613-907-7884

s.69(1)(g) re (a)

Senior Counsel, Treasury Board Secretariat Legal Services
Department of Justice Canada / Government of Canada
nathalie.sarault@tbs-sct.gc.ca / Tel: 613-907-7884 / TTY: 613-369-9371

Avocate-conseil, Service juridique du Secrétariat du Conseil du Trésor
Ministère de la Justice Canada / Gouvernement du Canada
nathalie.sarault@tbs-sct.gc.ca / Tél. : 613-907-7884 / ATS : 613-369-9371

This communication contains information that may be confidential, exempt from disclosure, subject to litigation privilege or protected by the privilege that exists between lawyers or notaries and their clients. If you are not the intended recipient, you should not read, rely on, retain, or distribute it. Please delete or otherwise destroy this communication and all copies of it immediately, and contact the sender at 613-907-7884 or by email at nathalie.sarault@tbs-sct.gc.ca. Thank you.

Ce message contient des renseignements qui pourraient être confidentiels, soustraits à la communication, ou protégés par le privilège relatif au litige ou par le secret professionnel liant l'avocat ou le notaire à son client. S'il ne vous est pas destiné, vous êtes priés de ne pas le lire, l'utiliser, le conserver ou le diffuser. Veuillez sans tarder le supprimer et en détruire toute copie, et communiquer avec l'expéditeur au 613-907-7884 ou à nathalie.sarault@tbs-sct.gc.ca. Merci.

**Pages 29 to / à 41
are withheld pursuant to section
sont retenues en vertu de l'article**

69(1)(g) re (e)

**of the Access to Information Act
de la Loi sur l'accès à l'information**

Standard on Automated Decision-Making

VERSION DRAFT IN DEVELOPMENT - v.0.5

<u>Version</u>	<u>Date</u>	<u>Updates</u>
0.01	March 1, 2018	Initial version - review by TBS CIOB and P&P policy experts
0.1	March 9, 2018	Version out to institutions - woohoo!
0.15	March 15, 2018	Version open for stakeholder contribution
0.16	March 17, 2018	<ul style="list-style-type: none"> • New 6.1.1.3 forbidding Very High Concern systems from total automation (unless legislation permits otherwise) • AIA change: Impact on liberty now an automatic High Concern or more
0.17	March 19, 2018	<ul style="list-style-type: none"> • New sections on procurement and existing systems • Revised introduction language and expected outcomes
0.18	March 26, 2018	<ul style="list-style-type: none"> • “Prevent unintentional bias” section under transparency changed to a new requirement under system auditing instead, requiring that institutions monitor for unanticipated outcomes.
0.19	April 13, 2018	<ul style="list-style-type: none"> • 3.5 revised to better reflect that Ministers were always accountable for their decisions • New 6.1 scope section - should clarify precision around what this covers. • 6.2.1 now requires the AIA before the DSS is used on a client. • Deleting third-party certifications for this iteration. There are no Canadian or international standards for which to certify against. • Added a 15 business day window for the GC CIO to object to the AIA, to ensure that departments are responded to promptly.
0.2	April 17, 2018	<ul style="list-style-type: none"> • Section on Benefits for Canada removed - Ministers have the prerogative to automate for whatever purposes they wish. • Old System Auditing section changed to Monitoring and Reporting; old M+R section folded into Transparency
0.3	April 19, 2018	<ul style="list-style-type: none"> • Roles and responsibilities dropped or integrated into requirements • AIA itself removed from the standard (still referenced, but will be a companion piece) • Restructuring of System Design section to increase simplicity in reading • Revised Expected Results • Scope tightened • Substantial amendments/precision to transparency requirements (e.g. new training requirement)

		<ul style="list-style-type: none"> • Formatting cleanup
0.4	May 4, 2018	<ul style="list-style-type: none"> • New title! May the fourth be with you • Advisory Board removed in lieu of a new peer review section • Requirement to issue open code included
0.4.1	May 15, 2018	<ul style="list-style-type: none"> • Definitions added • Scope statement has been narrowed and removed from requirements into its own section • Coming into force clause rolled into effective date
0.5	June 6, 2018	<ul style="list-style-type: none"> • New appendices for scaling requirements to level of impact • Moscope statement • Legal authority section amended • Amended contracting language, including IP clause better conforming to the <i>Policy on Title to Intellectual Property Arising Under Crown Procurement Contracts</i> • Automated Decision System now being used as the core term

Introduction

Promotes the use of automated decision-making while ensuring that its are compatible with core administrative law principles such as *transparency, accountability, legality and procedural fairness*.

s.69(1)(g) re (a)

1. Effective Date

- 1.1. This standard takes effect on (([REDACTED]))
- 1.2. All Decisions Support Systems that were in production prior to the coming into force of this standard, must complete an Algorithmic Impact Assessment and comply with all applicable provisions of this Standard within six months.

2. Application

- 2.1. This Standard applies to all institutions referenced in the ((*Policy under which this resides*)), unless excluded by specific acts, regulations or orders-in-council;
- 2.2. Agencies and Crown Corporations may enter into Specific Agreements with the Treasury Board of Canada Secretariat to adopt the requirements of this Standard and apply them to their organization, as required.

3. Context

- 3.1. The Government of Canada is increasingly looking to utilise technology and automated systems to make, or assist in making, administrative decisions to improve service delivery;
- 3.2. The Government of Canada is committed to ensure that the use of Automated Decision Systems are used ethically, and are compatible with core administrative law principles such as *transparency, accountability, legality and procedural fairness*;
- 3.3. This Standard is issued under the authority of section 7 of the *Financial Administration Act*;
- 3.4. This Standard supports the *Policy on Management of Information Technology*, *Policy on Information Management*, the *Policy on Service*, the *Policy on Privacy Protection*, and the *Policy on Government Security*;

4. Definitions

- 4.1. Definitions to be used in the interpretation of this standard are listed in Appendix A.

5. Statement

5.1. Objective

- 5.1.1. To ensure that Automated Decision Systems are deployed in a manner that minimizes risks to Canadians and federal institutions, and leads to more efficient, accurate, consistent, and interpretable decisions made pursuant to Canadian law and core principles of administrative law.

5.2. Expected Results

- 5.2.1. Administrative decisions are more transparent and accountable;
- 5.2.2. An increase in the use of automated systems to make, or assist in making, administrative decisions.

6. Scope

- 6.1. This Standard applies only to systems that provide recommendations to an authorized human administrative decision-maker ("Automated Decision System"). This includes systems that:
- 6.1.1. Classifies cases in terms of risk and priority;
- 6.1.2. Identifies cases for human review or investigation;
- 6.1.3. Provides overall recommendations about whether an application should be approved;
- 6.1.4. Renders the final administrative decisions.
- 6.2. This Standard applies only to systems that provide external services as defined in the Policy on Service.

7. Requirements

The institution's Chief Information Officer, as well as the Assistant Deputy Minister or equivalent are responsible for the following activities described in this section:

7.1. Algorithmic Impact Assessment

- 7.1.1. Complete an Algorithmic Impact Assessment, prior to the production of any Automated Decision System.
- 7.1.2. Apply the relevant requirements prescribed in Appendix C as recommended by the Algorithmic Impact Assessment.
- 7.1.3. Ensure that the Algorithmic Impact Assessment remains up to date and accurately reflects the functionality of the Automated Decision System.
- 7.1.4. Release the final results of Algorithmic Impact Assessments in an accessible format via Government of Canada websites and services designated by the Treasury Board of Canada Secretariat pursuant to the Directive on Open Government.

7.2. Transparency

Providing Notice Before Decisions

- 7.2.1. Provide notice to affected individuals at the earliest stages of administrative process that the decision rendered will be undertaken in whole or in part by a Automated Decision System, unless stated otherwise in legislation or regulation.
- 7.2.2. Ensure that any affected individuals can have access to information about the Automated Decision System's functionality. This includes, at minimum:
 - a. The role that the Automated Decision System has within the decision-making process,
 - b. A description of the training data, or a link to the anonymized training data if this data is publicly available, and
 - c. A description of the criteria used for making the decision, including business rules.

Providing Explanations After Decisions

- 7.2.3. Provide a meaningful explanation to affected individuals of how and why the decision was made as prescribed in Appendix C.

Open Source Code

- 7.2.4. Make available to the public all of the source code used for the Automated Decision Systems on a website or service designated by the Treasury Board Secretariat.
- 7.2.5. In cases where it is deemed that source code should not be disclosed, seek the approval of the Enterprise Architecture Review Board to exempt the disclosure. In these cases, the justification as to why code was not disclosed shall be published according to the process specified in the Directive on Open Government.
- 7.2.6. Ensure that all licenses required for the Automated Decision Systems are open licenses as listed in the Open Source Software Registry. In all cases, Canada must maintain the right to have access to foreground intellectual property to respond to any legal challenges.

7.3. Quality Assurance

Testing and Monitoring Outcomes

- 7.3.1. Before going into production, develop the appropriate processes to ensure that training data is tested for unintended data biases and other factors that may unfairly impact the outcomes.
- 7.3.2. Monitor the outcomes of Automated Decision Systems on an ongoing basis to safeguard against unintentional outcomes and to ensure compliance with institutional and program legislation, as well as this Standard.

Data Quality

- 7.3.3. Ensure that data being used by the Automated Decision System is routinely tested to ensure that it is still relevant, accurate and up-to-date and follow any applicable policy or guidelines with regards to data management practices.

Peer Review

- 7.3.4. Retain the appropriate expert to review the Automated Decision system, as provided in Annex C based on the Impact Assessment Level.

Training

- 7.3.5. Ensure that some employees are sufficiently trained in the design, function, and implementation of the Automated Decision System to be able to review, explain and oversee automated decision-making, as required.

Contingency

- 7.3.6. Ensure that a contingency systems and/or processes are available should the Automated Decision System be unavailable for an extended period of time.

Security

- 7.3.7. Conduct risk assessments throughout the development of the system and ensure appropriate safeguards to be applied, as per the Policy on Government Security.

Legal

- 7.3.8. Consult with the institution's legal services unit, to ensure that the use of the Automated Decision System is authorized by law.

7.4. Recourse

- 7.4.1. Affected individuals must be provided with information with regards to the options that are available to them for recourse to challenge the automated decision.

7.5. Reporting

- 7.5.1. Information on the effectiveness and efficiency of the Automated Decision Systems will be published annually on websites and services designated by the Treasury Board of Canada.

- 7.5.2. When requested, information on the achievement of the expected results of the Automated Decision System and compliance with this Standard will be provided to the Treasury Board of Canada Secretariat.

8. Consequences

- 8.1. Failure to comply with this Standard will result in the need to provide additional information relating to the development and implementation of compliance strategies in their annual report to Parliament. This reporting will be in addition to other reporting requirements and will specifically address the compliance issues in question.

9. Roles and Responsibility of Treasury Board of Canada Secretariat

The Chief Technology Officer for the Government of Canada is responsible for:

- 9.1. Setting government-wide direction on artificial intelligence generally, as well as establishing guidance for Automated Decision Systems.
- 9.2. Developing and maintaining the Algorithmic Impact Assessment and any supporting documentation.
- 9.3. Communicating and engaging government-wide and with partners in other jurisdictions and sectors to develop common strategies, approaches, and processes to support the responsible use of Automated Decision Systems.
- 9.4. Reviewing this Standard every three years after its effective date.

10. References

- 10.1. ((Relevant Legislation))
Financial Administration Act
Access to Information Act
Privacy Act
Security of Information Act
- 10.2. ((Relevant Policy Instruments))
Policy on Access to Information
Policy on Service
Policy on Government Security
Policy on Information Management

Policy on Management of Information Technology
Policy on Privacy Protection
Directive on Open Government

11. Enquiries

For information on this policy instrument, please contact the Treasury Board of Canada
Secretariat Public Enquiries.

Appendix A - Definitions

Automated Decision System

An Automated Decision System includes any information technology designed to provide a specific recommendation to a human decision-maker on an administrative decision, or designed to make an administrative decision in lieu of a human decision maker.

Administrative Decision

Any decision that is made by a Minister, a Minister's delegate, a court, or an administrative tribunal, authorized by legislation or regulation that affects the rights and/or interests of others.

Algorithmic Impact Assessment

A framework to help institutions better understand and mitigate the risks associated with Automated Decision Systems and to provide the appropriate governance, oversight and reporting/audit requirements that best match the type of application being designed.

Source Code

Computer program in its original programming language, human readable, before translation into object code usually by a compiler or an interpreter. It consists of algorithms, computer instructions and may include developer's comments.

Appendix B - Impact Assessment Levels

Level	Description
I	<p>The decision has a little to no impact on the rights or interests of an individual, community, organization, society, or the environment.</p> <p>Erroneous decision could reasonably be expected to cause nil to minimal harm.</p>
II	<p>The decision has a moderate impact on the rights or interests of an individual, community, organization, society, or the environment.</p> <p>Compromise could reasonably be expected to cause minimal to moderate harm.</p>
III	<p>The decision has a high impact on the rights or interests of an individual, community, organization, society, or the environment.</p> <p>Compromise could reasonably be expected to cause moderate to serious harm.</p>
IV	<p>The decision has a very high impact on the rights or interests of an individual, community, organization, society, or the environment.</p> <p>Compromise could reasonably be expected to cause serious to catastrophic harm.</p>

Appendix C - Impact Level Requirements

Requirement	Level I	Level II	Level III	Level IV
Peer Review	None	<p>At least one of:</p> <p>Qualified expert from a federal, provincial, territorial or municipal government institution</p> <p>Qualified members of faculty of a post-secondary institution</p> <p>Qualified researchers from a relevant non-governmental organization</p> <p>Contracted third-party vendor with a related specialization</p> <p>Publishing specifications of the Automated Decision System in a peer-reviewed journal</p>	<p>At least one of:</p> <p>Qualified expert from a federal, provincial, territorial or municipal government institution</p> <p>Qualified members of faculty of a post-secondary institution</p> <p>Qualified researchers from a relevant non-governmental organization</p> <p>Contracted third-party vendor with a related specialization</p> <p>Publishing specifications of the Automated Decision System in a peer-reviewed journal</p>	<p>At least two of:</p> <p>Qualified experts from the National Research Council of Canada or Statistics Canada</p> <p>Qualified members of faculty of a post-secondary institution</p> <p>Qualified researchers from a relevant non-governmental organization</p> <p>Contracted third-party vendor with a related specialization</p> <p>OR:</p> <p>Publishing specifications of the Automated Decision System in a peer-reviewed journal</p>
Explanation Requirement for Recommendation (6.1.1 and 6.1.2)	None	None	Meaningful explanation provided upon request based on machine or human review.	Meaningful explanation, including the variables that contributed to the decision,

				<p>provided with the decision rendered.</p> <p>Explanation can be human or machine generated.</p>
<p>Explanation Requirement for Decisions (6.1.3 and 6.1.4)</p>	<p>An explanation provided upon request based on machine or human review. This could include a Frequently Asked Questions section of a website.</p>	<p>Meaningful explanation provided upon request based on machine or human review.</p>	<p>Meaningful explanation, including the variables that contributed to the decision, provided with the decision rendered.</p> <p>Explanation can be human or machine generated.</p>	<p>Meaningful explanation, including the variables that contributed to the decision, provided with the decision rendered.</p> <p>Explanation can be human or machine generated.</p>
<p>Approval Requirement</p>	<p>None</p>	<p>Enterprise Architecture Review Board</p>	<p>Enterprise Architecture Review Board</p> <p>AND</p> <p>GC CIO</p>	<p>Requires specific authority from Cabinet</p>

Pages 55 to / à 69
are withheld pursuant to section
sont retenues en vertu de l'article

23

of the Access to Information Act
de la Loi sur l'accès à l'information

Responsible Artificial Intelligence in the Government of Canada

Digital Disruption White Paper Series

Version 2.0
2018-04-10

Responsible Artificial Intelligence in the Government of Canada

1. Table of Contents

1. Table of Contents	1
2. Version History	2
3. Executive Summary	3
4. Introduction	4
4.1. Objective of this paper and intended audience	6
4.2. Automation and Artificial Intelligence	7
4.3. Narrow and General Intelligence	10
5. AI for Smarter Government	10
5.1. AI for the Delivery of Services to the Public	10
5.1.1. Smarter Search	11
5.1.2. Chatbots	12
5.1.2.1. User Experience Considerations	14
5.1.3. Automated Decision Support	15
5.1.3.1. Appropriateness of Automation	16
5.1.3.2. Transparency and Recourse	17
5.2. AI to help design policy and respond to risk	18
5.3. Applying AI to the internal services of government	19
5.3.1. Information Management	19
5.3.2. Automated Content Generation	19
5.3.3. People Management	20
5.3.4. Security and Access Management	20
6. Policy, Ethical, and Legal Considerations of AI	21
6.1. Ensuring High-Quality Data	22
6.1.1. Prevention of Data Bias	22
6.1.2. Data for Insights and Privacy Rights	24
6.2. Transparency and Accountability	25
6.2.1. Accounting for the Actions of AI: The “Black box” Problem	25
6.2.2. Model Design and Outcome Biases	Error! Bookmark not defined.
6.2.3. Social Acceptability	26
6.3. AI and the Law: An Emerging Landscape	26
6.4. Technical Considerations	27
6.4.1. Cybersecurity considerations	27
7. Rethinking a Post-AI Enterprise	28
7.1. New Approaches to the Workforce	28
7.2. Evolving How Government Works	29
8. Conclusion	30

Responsible Artificial Intelligence in the Government of Canada

2. Version History

#	Date	History
Sections 0.1 – 0.10 –Early Concept		
0.1	July 21	Release
0.11	August 7	Release
0.12	August 17	Release
0.2	September 22	Release
0.3	October 2	Release
0.4	October 16	Release
Sections 1.0 – 1.X – Working Drafts for Broad Consultation		
1.0	October 27	First draft for open comment
1.1	November 6	New introduction, new section on security and user access control
1.2	November 21	New sections on inclusion and cybersecurity, revisions to sections on AI for policy, revisions to “evolving how government works,” new box on anthropomorphism, risk test appendix removed - it belongs more in directive format than in this white paper.
1.3	December 7	Version shared with FNIGC, several companies
1.4	March 5	Version sent to the Privacy Commissioner
Sections 2.0 - Senior Management Consideration		
2.0	April 17	Version for formal translation

3. Message from the Chief Information Officer of the Government of Canada

(to come)

4. Executive Summary

Artificial intelligence (AI) is a term used to describe a suite of related technologies intended to simulate and enhance human cognitive capabilities, such as pattern recognition, judgement, vision, or hearing. Having first been conceived in the 1940's, AI has advanced rapidly in recent years due to a combination of vast quantities of data, new mathematical techniques, and inexpensive computing power. AI systems now underpin many of the consumer products that Canadian use on a daily basis, from curating what media content we consume based on our interests, to helping us navigate our towns and cities. There are real-world examples of AI systems operating vehicles, writing newspaper articles, or generating art, that challenge previous assumptions of the types of tasks that can be delegated to machines.

Just as AI systems are rapidly transforming the world around us, so too is it expected that AI will transform the way that government operates. Imagine virtual service agents assisting Canadians and businesses with completing routine transactions 24 hours a day, seven days a week. AI systems can monitor the status of industries to detect early warning of regulatory non-compliance. They can sift through, structure, and recombine vast stores of data to help government institutions understand the information that they currently have, in order to more intelligently design public policy. These technologies have the potential to guide the public service towards a future of greater effectiveness and responsiveness to the needs of society than was ever possible before.

While the power that AI systems may bring to government could be significant, they must be deployed in a responsible and ethical manner. AI systems often require "training" using datasets that are reflective of the problem needing to be solved. If these data were collected or tabulated in a way that carries bias, then the outcome will be AI recommendations or decisions that are biased as well. Further, some AI systems currently operate as "black boxes," meaning that the decisions they make are difficult to audit or fully comprehend. In light of these limitations, it is important to understand where it is appropriate to deploy different types of AI systems, balancing the potential for gains in efficiency and effectiveness of government with the risk of misuse. Finally, although AI will afford institutions with new capabilities, institutions will need to apply a strong ethical lens to whether the technology should be deployed at all in certain circumstances.

Responsible Artificial Intelligence in the Government of Canada

AI is a capability that rests atop an expert and disciplined data science practice within institutions, as well as leveraging Canada's leading AI talent base. These systems will challenge how government institutions work, demanding a prioritization of good data governance practices, and requiring new skillsets of knowledge workers.

This paper proposes a set of seven principles that will be expressed in all future Treasury Board policy on the use of AI systems in government:

1. People should always be governed – and perceive to be governed – by people;
2. AI systems deployed on behalf of government should be trained to reflect the *Values and Ethics of the Public Sector* as well as Canadian and international human rights obligations; they should be used to reinforce these values where possible;
3. Organizations are accountable for the actions of AI systems, and should build systems that are auditable;
4. Understanding the need to protect privacy and national security, AI systems should be deployed in the most transparent manner possible;
5. Organizations should ensure that reliable contingencies are in place for when AI systems fail, or to provide services to those unable to access these systems;
6. AI systems should be developed in a diverse team that includes individuals capable of assessing the ethical and socioeconomic implications of the system;
7. AI systems should be deployed in a manner that minimizes negative impact to employees where possible, and should, where feasible, be created alongside the employees that will work with them.

5. Introduction

First it was chess, then Go, then poker. One by one, we have taught machines to exceed us in some of our most treasured – and complicated – games. These accomplishments showcased advancements in techniques achieved much faster than predicted, and were at least partially responsible for kicking off an era of massive investments and excitement in artificial intelligence. We have trained machines to mimic the outcomes of human learning and decision processes, such as adaptation, bargaining, and bluffing. With successive and public displays of computing prowess by the likes of IBM, Deepmind, or Facebook, and the rapid growth of a startup ecosystem, advances in AI have begun to dominate the press and capture the public's imagination.

While AI was originally conceived in the 1940's, over the past decade, these applications have been deployed in such variable and extensive ways that it increasingly drives the modern economy. AI has

Responsible Artificial Intelligence in the Government of Canada

replaced humans on stock market floors¹ and in the management of multi-billion dollar hedge funds.² It assists with medical diagnoses and operates complex machinery autonomously. It has been applied to corporate process and workflow automation to increase efficiency of their operations. AI agents are beginning to use natural language effectively enough to interact with humans via intelligent chatbots. There is a very high likelihood that by 2025, AI will touch every aspect of modern society in ways both visible and invisible to Canadians.³

Since the 1970s, early investments in Canadian researchers allowed an AI industry to bloom here. The advances of Canadian pioneers in machine learning positioned this country as a global leader in AI research, development, and application. Budget 2017 committed \$125 million to launch a Pan-Canadian Artificial Intelligence Strategy to support these clusters and attract the talent they need to maintain their advantage. Establishment of superclusters in Montreal, Toronto, and Edmonton has seen both the rise of world-leading research institutes as well as an ecosystem of AI startups that are internationally competitive and driving innovation.

Now, the Government of Canada is looking into how it can harness the opportunities provided by AI to offer novel and more timely services to citizens and other users,⁴ as well as improve the effectiveness and efficiency of its operations. Federal institutions are working towards offering better user experiences to make their services easier to use, but these gains will not accomplish a frictionless service environment if the person faces weeks-long backlogs in having a benefit application processed. Especially in circumstances where work is routine, AI systems can work faster and often more consistently than humans performing the equivalent tasks, and will work over evenings, weekends, and statutory holidays. Their capacities for decision-making are not adversely affected by physical fatigue or the natural emotional and relational situations people face based on their natural makeup. AI systems can be deployed by service institutions to answer questions posed by users – as well as make eligibility determinations – in order to dramatically improve the response time of service.

On the other hand, when administrative tasks are complex and value-laden, it can be difficult to ensure that the actions of the AI systems align with the spirit and intentions of the policy being implemented. Working with complex social and economic systems is considerably more complex than a game of Go. How do we know whether an AI system is appropriately trained for its task, and that data is interpreted in a manner that is accurate and responsible? How do we know whether AI is making biased or prejudicial decisions? How can AI systems be coded to meet similar legal obligations as human public servants, such as the Charter of Rights and Freedoms or the *Privacy Act*, and who is responsible when they fail to meet

¹ See example: <http://www.bbc.com/news/business-34264380>

² See example: <https://www.theguardian.com/technology/2016/dec/22/bridgewater-associates-ai-artificial-intelligence-management>

³ A qualitative survey by the Pew Research Center of over 2,500 academics, policy analysts and corporate executives found broad consensus to support this prediction. While the study was American, respondents were international. See: Pew Research Center, “AI, Robotics and the Future of Jobs.” Link: <http://www.pewinternet.org/files/2014/08/Future-of-AI-Robotics-and-Jobs.pdf>

⁴ This paper uses the term “users” to represent the diverse groups that use Government of Canada services including, but not limited to, citizens, permanent and temporary residents, and businesses. It avoids the term “client” to reduce confusion with the legal term.

Responsible Artificial Intelligence in the Government of Canada

these obligations? How do we teach it social, cultural, or geographical context such that it can make decisions in a nuanced fashion? How do we know the rationale behind the decisions of an AI system? What types of decisions should always require some form of human intervention? How do we know that the data on which an AI system is trained, which is sampled from real data about real Canadians, is kept secure and private once the AI system is in deployment? What are the workforce requirements in a post-AI world?

Governments worldwide are now grappling with the consequences of a technological development that is transforming service delivery across sectors. The United States, United Kingdom, France, the United Arab Emirates, China and Japan are just some of the jurisdictions that have undertaken high-level examinations of AI systems within their respective governments and on their economies writ-large. The Government of Canada has the opportunity to build on the brain trust of private sector and academic leaders in this field to position itself as a world leader in AI for policy development and service delivery. It has the opportunity to signal to all sectors that AI can be harnessed in a manner that is ethical and supportive of positive outcomes for Canadians without sacrificing the benefits of the technology.

While AI is undergoing rapid advancement, it is important that the policy, ethical and legal implications of the use of this technology to deliver government services be addressed methodically and with an understanding of this complexity. The service delivery opportunities are significant, as are the pitfalls.

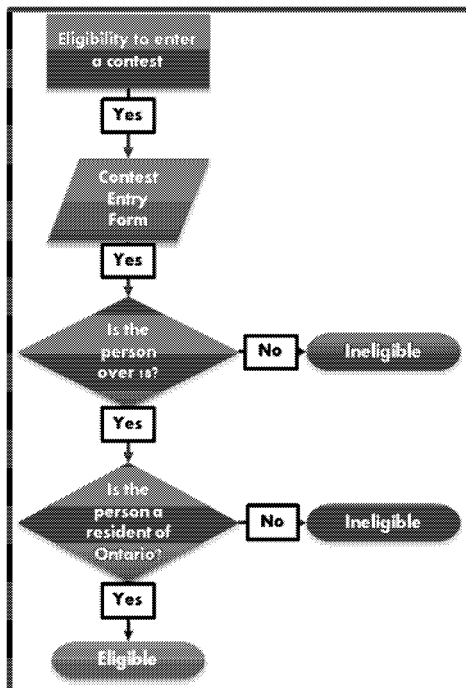
5.1. Objective of this paper and intended audience

The scope of this paper is limited to the specific use of AI applications by federal institutions for their own use only; it does not touch on the Government's response to automation in the private sector and its effect on society. This scope is broadly aligned with the mandate of the Treasury Board in its role in setting general administrative policy for federal institutions.

This white paper will examine the policy, ethical, technical, and legal considerations around the use of this technology within the Government of Canada. Its primary objective is to assist federal institutions by providing recommendations on how these systems should be implemented. The intended audience is therefore broad, from Deputy Heads or Chief Information Officers wishing to understand a significant new technology, to policy managers or service designers looking to apply AI to the programs or services that they provide. At the same time, it is intended to communicate to the AI development ecosystem in the academic and private sectors the use cases and policy considerations that are common in the federal government.

Throughout the paper, illustrative examples are used to show how this technology can be beneficial to users. Unless otherwise specified, these examples do not represent any existing plans of the Government of Canada and should be considered theoretical only.

5.2. Automation and Artificial Intelligence



Humans have always been intrepid designers of tools. From the scythe and wheel to the internal combustion engine and the computer, we have always designed tools to produce more from less. For most of human history this has led to technologies that have extended our physical capacities, but with the outbreak of the Second World War, humanity started designing tools that started to extend our cognitive and analytical capacities as well, such as memory, attention, judgement and decision-making. In a sense, we started designing brains for our tools.

We eventually designed tools that took over tasks for us completely. Automation has been a hallmark of industrialization since the robot Unimate was deployed in a New Jersey GM plant in 1961 for hazardous die casting, not just for physical tasks, but for analytical ones as well.

Behind the automated processes that drive the 21st century economy are a series of logical instructions known as algorithms.

Like a recipe, algorithms are processes that inform a machine how to perform a specific task. They can often be broken down into a series of decisions that are defined by the programmer; such as “is the individual over 18 years old?” or “is the individual a legal resident of Ontario?” The output is decided based on these decisions. The rules of these algorithms do not change unless programmers decide to change them. Closed-rule algorithms are used in the support of decisions widely in the private and public sectors today; for example, the Canada Revenue Agency uses closed-rule algorithms to support tax processing, with the rules defined by legislation and regulation.

Enter Artificial Intelligence

While it was the eminent British computer scientist Alan Turing that first conceived of “the thinking machine,” the term “artificial intelligence” was coined later in 1956 by the American computer scientist John McCarthy to describe “the science and engineering of making intelligent machines.” As technology has evolved, AI has grown to become a term that includes a broad spectrum of related technologies that seek to imitate and enhance aspects of human intelligence, such as vision, identifying patterns in information, or understanding language. In a sense, AI is when computers do what only humans could before. The term is used to describe applications as innocuous as a system that recommends books to read, to fictional advanced human-like intelligence capable of everything a human is. As such, there is no single, internationally-recognized definition for AI, and the term may mean different things to different people.

The development of **machine learning** was a critical milestone. Machine learning is a method by which algorithms can be trained how to recognize patterns within information, and the ways in which data interrelate. For example, a learning algorithm that recommends books based on your purchasing history

Responsible Artificial Intelligence in the Government of Canada

provides better recommendations as you purchase more books. It does this without a human on the backend needing to adjust the programming instructions. If that algorithm had access to your browsing history as input data - and assuming that it was programmed to know what to do with that data - its recommendations may improve even more because it begins to “know” your tastes better.

Machine learning is by no means the only application of artificial intelligence. **Natural language processing** allows computers to parse meaning and context out of written text. This is used extensively, for example, in legal analysis software to derive insights from large volumes of text. **Machine vision and hearing** provide machines with the capability of structuring, and using, typically unstructured data such as imagery or sound. This is used in a diverse range applications, from autonomous cars “seeing” obstacles to smartphone applications that can identify a song played in public.

Either one or a combination of these techniques underpin many of the private sector digital services that people use regularly worldwide. Major social networking platforms, media platforms, and smartphones all run machine learning algorithms that provide services such as navigating traffic or curating news. It is not necessary to use machine learning in all approaches to automation; for applications where rules are precisely defined (such as the example above), a closed-rule algorithm is sufficient for the task.

Early experiments have existed since the late-1950s to show how machines are capable of learning and self-improvement. Today, researchers and developers have access to powerful and inexpensive cloud computing resources, parallel computing, as well as profoundly more data. Smartphones and the sensors located within them, coupled with the popularity of social media and internet culture means that a typical person produces a bounty of harvestable data every day - even when they are sleeping.⁵ As a result, the development - and implementation - of AI has progressed rapidly in the last ten years. As the Internet of Things connects common consumer products and appliances to the internet, the data points that we generate in our day-to-day lives will likely grow exponentially.

This ability to capture and use data in unprecedented ways has had a direct impact on the development of AI because of these technologies’ need for sufficient quality and quantity of data. Think of AI as a very sophisticated engine; without data to fuel it, it can’t propel the vehicle. Data needs to be available in sufficient quantity, they need to be relevant enough to the task at hand, they need to have been collected and described in a manner that is free of bias, and they need to be in a format that is readable by a machine. Despite addressing AI, much of this paper is devoted to issues surrounding data rather than the instructions precisely because insufficient quality and quantity of data can render the most expertly-programmed AI useless - or worse - harmful.

We are now at a point where machine learning can enable AI not only to replicate many human tasks - it can come close to surpassing our effectiveness at certain tasks, such as recognizing subjects of images,⁶ or reading lips.⁷

⁵ For example, by using an app that monitors sleep time and quality.

<http://ns.umich.edu/new/multimedia/videos/23822-smartphones-uncover-how-the-world-sleeps>

⁶ Based on 2017 results of the University of Washington MegaFace challenge:

<http://megaface.cs.washington.edu/results/facescrub.html>

⁷ Based on LipNet results. See: <https://www.technologyreview.com/s/602949/ai-has-beaten-humans-at-lip-reading/>

Responsible Artificial Intelligence in the Government of Canada

Advances in techniques

There are many approaches that developers take to AI; for example, **deep learning**, a branch of machine learning, has been used extensively in modern private sector services. While many deep learning algorithms use labelled data, it also brought the capability of using unstructured data such as audio or visual data, allowing the system to extract features of information on its own.

There have been significant advances in **artificial neural networks** in recent years. Inspired by the human brain, neural networks are composed of artificial neurons, which receive data individually and calculate outputs independently, allowing a complex problem to be broken down into millions of simple problems and then reassembled as one answer. As the network is provided more data, it can identify new and complex relationships in data, much like how the human brain forms synapses. This complex relationship is encoded in the weights, learned during model training, that connect the neurons in the neural network.

For example, rather than just learning what a bear is based on analyzing millions of images tagged as “bear,” a deep learning AI can extract features from images of a bear on its own. Humans do that as well; we learn a bear’s size and shape, where a bear may be found, typical colours of its fur and its family structure. That way, when we see an image of a bear that we have never seen before, we can infer that what we are seeing is a bear based on understanding its components.

The complication of deep learning is that it is not always possible to have access to massive data and to understand the importance associated with different variables of the problem. Using the above example, it is very difficult to understand whether an AI neural network considers fins as important than scales in determining whether something is a bear or not, both because the network is complex, but also because as the network is exposed to more examples of bears, this weighting may change. This process is often reliant on very large volumes of data that are broadly representative of the world within which the system will operate; for example, an autonomous vehicle trained exclusively in the UK could not be deployed in Canada, where driving is on the opposite side and some rules differ.

Another approach is, reinforcement learning; this is a subset of machine learning whereby machines are trained by being rewarded for desired outcomes and punished for undesired ones, similar to how we train dogs to play fetch. Rules are provided to the algorithm as to what it must do to earn a reward; for example, if the bear brings the ball back, it will get a fish to eat. The bear will not receive a fish if it does not bring the ball back. Reinforcement learning is especially useful in situations with well-defined outcomes, for example games and puzzles.

Reinforcement learning algorithms can be trained in advance using simulations, but they adapt more quickly once able to interact with its intended operating environment. However, they need clear definitions of “right and wrong” - outcomes that are desirable or undesirable, and the choices of those definitions are laden with values.

The choice of methodology will matter depending on the problem needing to be solved.

Responsible Artificial Intelligence in the Government of Canada

5.3. Narrow and General Intelligence

Whereas you are a multifaceted individual with a number of potentially unrelated interests, AI is often targeted for a single objective or task. This is known as “narrow” intelligence; while it can excel at one task – even surpassing a human – it cannot learn a second task without being explicitly targeted to do so. For example, while you may be a software engineer that speaks four languages fluently and is an amateur chef, an AI system trained to identify high-risk travellers cannot simply choose to learn to translate languages. This is because AI is software and does not have agency.

While research is underway to determine whether AI can achieve general intelligence, this achievement is still highly theoretical. A generally intelligent AI brings with it significant policy implications as well, but this paper will focus on the implications of narrow AI.

AI is software, not an organism

For decades, science fiction has introduced AI characters – whether in robot or incorporeal form – to the social consciousness. The popularity of characters like HAL 9000 or C-3PO may cause us to ascribe some degree of personification to AI. While it is designed to mimic human intelligence, the “learning” and “understanding” that a machine undergoes is different than the biological processes that we humans rely upon.

This paper refers to AI using humanlike semantics from time to time because it is a helpful way to communicate technical concepts, but it is important to remember that fundamentally, AI is software, not a conscious being, and should not be ascribed agency over its actions. Doing so removes the accountability of an organization over its software.

6. AI for Smarter Government

AI is not a technology looking for a problem; it is a suite of tools with the potential to help the GC deliver services more effectively, design policy more responsively, and potentially enable an entire suite of new capabilities in designing policy and delivering services. As the set of applications is diverse, its potential impact on the public sector is wide-ranging. Institutions have been examining applications that can be organized into three interdependent themes:

1. Applying AI to the delivery of services to the public
2. Applying AI to help design policy and respond to risk
3. Applying AI to the internal services of government

6.1. AI for the Delivery of Services to the Public

End-to-end digital self-service is the norm throughout much of the private sector service spectrum. The ability to access the entire continuum of the service from application to delivery without the need for a paper form, or for the user to have to interact with a service agent, is typical. Ideally, the service experience from authentication to application to receipt of benefit or issuance of payment should be a

Responsible Artificial Intelligence in the Government of Canada

seamless process that does not require the use of a phone or visit to a service centre unless chosen as the preferred way to receive service.

The government has decided to prioritize the development of digital services. Phone and in-person channels are inherently less convenient for users, as opening times are restricted, require waiting on hold or in line, or involve travel times. For individuals and businesses alike, lengthy wait times, or the requirement to access services during business hours can lead to an unacceptable loss of leisure time or productivity. Assuming that the digital service offering is understandable, convenient, and accessible enough for someone to want to use it, there is incentive for all parties of a service transaction to want to move to the digital channel. According to the Canadian Radio-Telecommunications Commission, broadband access in Canada will likely reach 90% as soon as 2021, and digital services will be more within reach for the vast majority of Canadians.⁸

Even if all services were provided digitally, as of today, there may be services by which some will elect to use other channels. There are some complex or sensitive needs that may demand more nuanced or personal service provision. Some people simply may feel more comfortable raising their issues in front of another person. In these circumstances, people may use an alternative channel such as phone or in-person if these are accessible to the individual. Even in these cases, AI can empower services by providing faster decisions, or tools that provide an overview of the individual's sentiment during the progress of the call.

More intelligent digital tools interacting directly with a user can play a role in keeping them on the digital channel. Smarter search and chatbots are capable of parsing natural language into searchable terms, accessing information located in FAQs, manuals or even specifically-identified internal documents and reply to the question in a way the user can understand. With additional information and user feedback, these tools will continuously improve at this task without the need for direct human intervention.

6.1.1. Smarter Search

Building a website targeted at millions of people presents a challenge; people interpret information differently, and may have different expectations as to where information can be found. Usability testing can help understand how people are interpreting information on a website, but advances in natural language processing (NLP) make the task of finding relevant information much easier than it used to be.

NLP technology parses natural language into underlying meaning, which then can be used in service of some task. For example, if a user loses their job, rather than having to look up Employment Insurance specifically, search for "I've lost my job" and see results that are relevant to that request. Over time, the application learns the relevance between search statements and the services that people are looking for. This is superior to older search methodologies, which would literally scan for the statement "I've lost my job" in web content. Over time, the algorithm will learn more patterns and do a better job at understanding what users want. NLP search functionality is widely used in the private sector today.

⁸ CRTC <http://www.crtc.gc.ca/eng/internet/internet.htm>

Responsible Artificial Intelligence in the Government of Canada

6.1.2. Chatbots

Chatbots are virtual user service representatives that offer capabilities of searching for information, or escorting a user to the right webpage. They work similarly to NLP search, but add a layer of interactivity and personalization.

The capabilities of an AI chatbot can be scaled up over time to provide expansive levels of user care as it gains experience and improves the way it manages information. It can offer responsive services, answering queries related to services passively. Eventually it can expand to become more navigational, offering hints, advice, or step-by-step instructions more reflexive of where a person is in the continuum of their service experience. Eventually an AI can be capable of actually executing instructions, such as accessing and pre-filling a form based on natural language. However unlike a website where a “what’s new” section can easily communicate new information or services available, some thought must be given to how the end user is aware of new functionality of a chatbot.

A chatbot may be offered to clients embedded in your webpage, or within another platform where your users are commonly found, such as SMS text messaging, Facebook Messenger, WhatsApp, Twitter, or Slack. This technology has advanced significantly over the past five years and is expected to continue its rapid advancement for the next decade, for providing both external and internal services.

Chatbots offer a diverse opportunity to provide services to users. Chatbots help filter routine questions away from human service agents so that they may focus on helping users through complex or distressing cases, or cases where a user is uncomfortable relaying their circumstances to a machine. They may also assist with public consultations on policies or programs, by being able to ask follow-up questions and react to user feedback in a much more nimble fashion than a survey.

This technology has been deployed successfully in the public and private sector. The United States Citizenship and Immigration Services uses a chatbot named Emma to answer users’ questions and provide a pre-check for eligibility. Emma not only answers questions, but provides navigational services; the search query “I’ve been offered a job in the US” not only provokes a response from Emma, but brings the user to the “Working in the United States” site. The bot is trained in English and Spanish. Another bot, Sgt. Star, is deployed by the US Army to answer questions to prospective recruits.

Institutions looking to deploy chatbots will need to ensure that there is training data available for the bot to learn the appropriate terminology for the service. This data can include previous interactions with clients looking for the service in question, whether emails, chat logs, transcripts from phone conversations, or social media. Ideally the datasets would include data on the outcome of the service interaction as well to ensure that responses to questions are those that actually satisfy clients.

Chatbots have limitations. As described above, conversations carry a lot of information outside of the basic text. Emotional queues or the use of sarcasm and humour can quickly confuse an AI conversational agent, or teach it bad behaviour. While they are adept at managing basic questions, a lengthy, interactive conversation is not possible at this time. Some chatbots provide a user with a defined set of potential inputs to reduce errors in the conversation, which results in a more scripted interaction. This can be useful

Responsible Artificial Intelligence in the Government of Canada

for quickly helping users find the information they need, although scripted interactions quickly become difficult to control as the scope of the bot's responsibilities increase.

An additional benefit of chatbots is their ability to structure data through a standardized approach to collection. Through interactions with users, a chatbot can help reduce spelling errors, inappropriate entry of dates and addresses, etc. This improves overall data quality, which in turn could help eligibility determination.

Chatbots offer transactional capability as well, merging the functions of both a virtual front-line service agent and the application form by collecting information directly from the user or their file in the institution's Customer Relation Management software.

It's important to remember that a user interacting with a chatbot may ask questions that are well outside the scope of its expertise. Users may disclose important personal information even when advised not to; they may even require immediate emergency assistance. In such a circumstance a human would be guided by a mix of their training and their own moral compass, but machine intelligences would need a means to triage these events, as well as pre-programmed responses.

Just like a human agent, a chatbot needs to be treated as an agent of the organization, which means that the information that it provides must always be accurate and up-to-date. Learning chatbots may provide advice to Canadians and, like humans, sometimes make mistakes. For example, a chatbot may give a person the wrong form or provide them an incorrect deadline. Chatbots that are designed to actually replace a form through conversational means may misinterpret input and submit incorrect information.

There have been significant and swift advances in chatbot technology, but despite these advances, it is a long way from flawless. In the future, bots have the potential for replacing forms as a way to collect information from users. They may even emerge to become the primary service delivery platform. Assuming that they have access to the widest range of information possible, bots can theoretically inform a user about any service in any institution with an almost expert like knowledge, far surpassing the ability of one individual's recall.

Finally, there are those in Canada who do not have access to reliable broadband internet, and may not in the near future. It is important that institutions continue to cater to these users and do not solely rely on chatbots for front-line services.

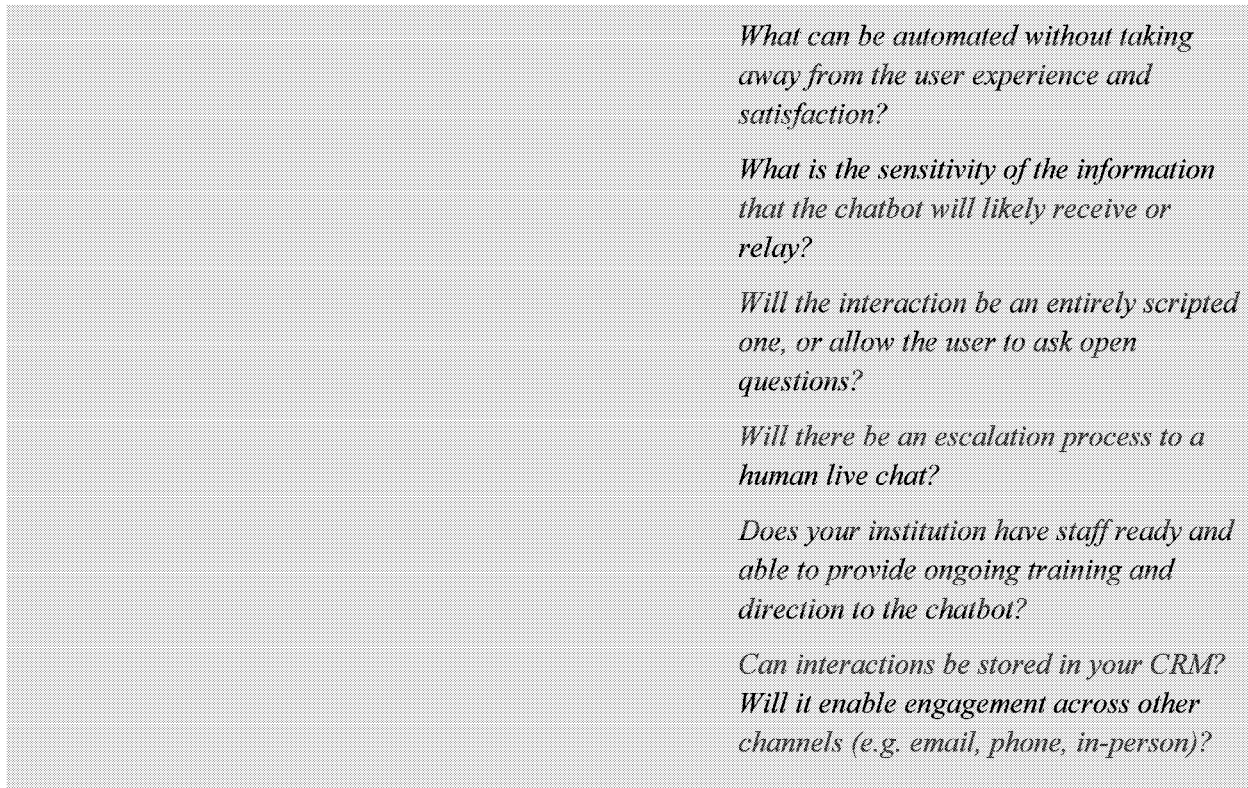
Is your institution ready for a chatbot?

When determining whether to deploy a chatbot, an institution should be able to answer the following questions:

Is there a clear business driver for the chatbot? Does your institution receive a high volume of routine inquiries?

Are the most common inquiries known and are data available to answer them?

Responsible Artificial Intelligence in the Government of Canada



6.1.2.1. User Experience Considerations

The GC has a wide policy and service landscape; if chatbots speaking to these policies and services offer interaction experiences that differ significantly, then users' acceptance of this technology can suffer and benefits will be unrealized.

A chatbot should not be used as a substitute for good discoverability of information on a website; it can add supplementary information or clarification to a user, but should not be seen as to replace the need for a well-designed site.

Chatbot conversations should be introduced with a brief privacy notice that is compliant with the *Treasury Board Standard on Privacy and Web Analytics*. This notice should provide a link to a page with more information on the information collected in the course of the conversation, including any metadata, for example: time and date, duration, whether the conversation was ended by the user or the agent, whether and when the discussion was escalated to a human, etc. Additionally, users should be informed that they are communicating with a chatbot.

Bots should be able to relay information in a professional tone as a representative of the Government of Canada. Machine learning chatbots may learn language that is potentially unprofessional, abusive, or harassing if exposed to sufficient examples. Where possible, institutions should work with vendors to prevent them from learning this behaviour, whether using a keyword blacklist, or other methodology. It is important to be continually monitoring chatbots' performance in this regard.

Responsible Artificial Intelligence in the Government of Canada

Some institutions may choose to use avatar, which is a personification of the chatbot. Visual avatars that express some emotional range improve users' belief in the competence of the virtual agent.⁹ The question of whether or not a chatbot should be gendered as male or female - or, for that matter, anthropomorphized (meaning: made to appear human - deserves close attention. It is unclear whether the use of a female gendered "assistant," could serve to perpetuate false, misleading and ultimately harmful cultural stereotypes about the status of women. To avoid a misstep in this sensitive area, some organizations have made the proactive decision to characterize their assistants as androgynous, such as Capital One's *Eno* and Sage's *Pegg* or non-human, such as Google's *Voice Assistant*.¹⁰

Institutions should be mindful that people in rural or remote locations may encounter latency that will affect their ability to respond to the chatbot's queries. It's important to ensure that response times from the user are permissive.

Chatbots must be accessible and meet accessibility standards and requirements of the GC. It is also important that chatbots be able to be read by screen readers, or are able themselves to communicate vocally, for persons with visual disabilities.

They should use plain language so as to be understood by users with varying levels of education or comfort with Canada's official languages. There is an opportunity to offer chatbots in a wide variety of languages should enough training information be available. Users should be provided with a clear escape from the conversation. If a user finds that a chatbot is no longer useful, or is incapable of answering their query, there should be a clear means to transfer the conversation to a human agent (if available), or to send email correspondence. Additionally, if a chatbot has answered a query and the user has ended the session or refrained from answering another question, the chatbot should politely end the conversation.

6.1.3. Automated Decision Support

Improving users' experiences when interacting with government services is important, but the benefits of this work are lost if the wait time to receive eligibility decisions on services is too long. Part of service excellence is cutting wait times, and AI can play a role.

To start, AI can be applied to electronic forms – both user-facing and back-end – to help ensure that data entered meets your institution's standard of quality. This modest application can greatly assist your institution's ability to use the data for decision-making later on.

Processing service applications requires that an analyst review application information, verify to see if it is true and believable, and checking if the information that has been submitted meets the program's

⁹ [6] Demeure, Niewiadomski and Pelachaud, "How Is Believability of a Virtual Agent Related to Warmth, Competence, Personification, and Embodiment?" *Presence*, October 2011. Link: http://www.mitpressjournals.org/doi/pdf/10.1162/PRES_a_00065

¹⁰ For more on Eno and Pegg see: <https://www.accountingtoday.com/opinion/the-tech-take-the-genderless-face-of-accounting-bots>

For more on Google's Assistant, see: <https://www.engadget.com/2016/10/07/google-assistant-desexualize-ai/>

Responsible Artificial Intelligence in the Government of Canada

eligibility criteria. This process can take time, both due to the amount of information collected as well as the limitations on resources.

By using appropriate program-related input data and a model to test inputs against rules, such as legislative or regulatory requirements, an automated system may be able to process eligibility decisions faster than and as well as a human in many circumstances. This allows eligibility analysis to be processed outside of core work hours, for data analytics to be gleaned and acted upon promptly and organically, and for patterns to be established so that particularly complex or unexpected applications can be investigated more thoroughly. Strictly speaking, this approach can be done without the use of AI, as the rules themselves are strictly defined by the institution.

This level of decision automation has been tested and deployed in private sector settings for over a decade. Insurance and financial sectors have been pioneers in decision automation to improve service response times and to increase fraud detection. These sectors have similar challenges to governments: mission-critical systems with many dependencies, limited budgets and competing priorities for IT development, and a desire to maximize transaction throughput and minimize fraud.¹¹

What if the system was designed in such a way that humans did not choose the eligibility criteria at all, but allowed a machine to determine what applicants should be eligible based on desired outcomes? For example, imagine a hypothetical program that provides small grants to exporters. Rather than have the program experts select the eligibility requirements themselves, an AI system analyzes similar firms in similar industries, and determines the likelihood of success following the grant. Of course, choosing the metrics that define “success” remains the responsibility of the program, but the criteria may vary. Perhaps there are different predictors of success for different sectors, or predictors that human analysts missed.

This approach has the potential to provide services with more effective outcomes, but brings challenges. For example, criteria are often enshrined in legal authorities. If there is a challenge to the decision, the institution would require to show what criteria were used to make the decision, something that might be difficult to show using current technology. This issue is further elaborated below.

Many government services have existed for decades; assuming there is high-quality, machine-readable data available, there is a significant volume of potential training sets to train AI how to process eligibility. By showing AI examples of successful versus unsuccessful applications, it can determine the necessary patterns to extend this reasoning to a new application on its own, effectively mimicking the experience of a human. For this to work, institutions need to have data on the outcomes of services in a format that is readable by machine.

6.1.3.1. Appropriateness of Automation

Should a service be automated completely from end-to-end, or should human intervention and approval always be required? The suitability of an automated system to deliver end-to-end services must be

¹¹ See McKinsey report, “Automating the bank’s back office,” Link: <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/automating-the-banks-back-office>

Responsible Artificial Intelligence in the Government of Canada

analyzed on a case by case basis. Much depends on the type of decision being made and the amount of discretion that any particular decision requires. Departments will have to carefully consider:

- Whether they are acting within lawful boundaries;
- Whether additional authorities are required;
- The procedures and mechanisms to be implemented to ensure transparency and to be able to document how a decision was reached, especially when such a decision affects individual rights and privileges and involves the exercise of discretion.

A “human in the loop” may not be straightforward in overturning machine decisions. Unless they are specifically instructed to, human officers will need to bring themselves to question the authoritativeness of the machine recommendation. Enough information would have to be provided to the human - both from the original input data such as a benefit application - but also around the rationale behind the decision. The human analyst should be required to document why the machine recommendation was not followed. Machine decisions flagged for human approval or overturning would themselves have to be monitored to ensure that there is no internal conspiracy or mismanagement.

The Government of Canada provides a diverse set of programs and services across over 140 federal institutions. Some of these programs and services are critical to the fundamental well-being of people, the economy, and the state; others are less. Should the same rigorous governance and accountability measures be required for non-critical programs as critical ones? Can we classify programs and services into risk categories to better target governance to be proportional to risk?

As TBS prepares guidance on how institutions can responsibly introduce automated decision support to their organization, it will develop a tool by which institutions can assess the degree of automation that is appropriate for their program. Guidance on governance could then be linked to the risk score.

6.1.3.2. Transparency and Recourse

How much information should be provided to users on the decision-making process? The ability and need to explain algorithmic decision making requires a delicate balance. On one hand, transparency builds trust and social acceptance, and provides users with information with which they can challenge decisions and business processes. On the other hand, providing too much information to the public can open a door to malicious manipulation of the algorithm.

Users should be notified in advance of submitting an application that it will be processed by an algorithm, along with a link leading a webpage with accessible, non-technical information on the decision-making process. This information should include a description of the sources of data used to make the decision, and links to recent system performance audits.

Further research is required to determine whether users should be provided an opportunity to opt out of automated decision making in advance of applying for a service. On one hand, this provides users with more control over how their personal information is handled. On the other hand, designing systems for this to occur may be impractical and expensive. Regardless, in the event of a negative decision, users

Responsible Artificial Intelligence in the Government of Canada

should be provided with an opportunity to have their application revisited by an informed human case assessor.

Further research is also required on what information institutions should provide on the design and functionality of AI tools (algorithms, logic, decision making rules), understanding that algorithms may be manipulated with too much of this information.

Regardless of the methodology used, it's important that institutions only automate a process when they have obtained a high level of confidence in the decisions that it is making in a test environment.

6.2. AI to help design policy and respond to risk

What if we were more accurately able to predict migration flows, forest fires, or the impact of an aging population? What if we knew in advance which ports of entry would be more likely to encounter contraband, or which consumer products might be more susceptible to recall? Existing analytical models have already given the GC the ability to better understand certain social or environmental outcomes to policy, but with new methods able to identify patterns in data that perhaps humans were previously incapable of doing, we may be able to make more precise and informed predictions than ever before.

Governments work with big problems. We work in an environment often marked by complex, interdependent systems, where small policy changes can result in massive impacts among a population or the economy. If we can use data to predict the impact of our work with greater precision, or to understand future pressures on social or economic programs, then we can respond more efficiently and ensure that regulatory resources are focused on the highest risk elements of their industries.

Using both structured and unstructured data sources, institutions can enhance their ability to understand what is happening in society and the economy, both in Canada and beyond. This will allow for more effective regulation of industries, as well as more informed policy planning through the use of simulation. The ability to combine even anonymized data sets across institutions in real time may be able to provide policymakers with new insights as to what is causing certain outcomes in society.

There are some limitations to this approach. Predictions are extrapolations of patterns that appeared in the past; while access to vast data sets brings greater opportunity to predict in a complex system, AI can't make truly novel predictions, because the past is not necessarily an indicator of the future. Like all AI systems, the right quantity and quality of data will need to be accessible to make accurate predictions. There is also a risk that predictions are made using data that has been collected in a way that is biased or not fully representative of the world that we live in; this issue is further discussed below.

Already, many federal institutions use a method to describe and compare the degree of risk involved with providing a service to a user. This "risk scoring" technique can be an efficient method to associate an administrative action with risk. To date, this has most often been accomplished using methods that require institutions precisely defining what risk is in their universe. These "closed-rule" algorithms, while not AI, are a form of automation that has shown to be service-enabling by reducing compliance and enforcement burden on lower-risk users.

Responsible Artificial Intelligence in the Government of Canada

6.3. Applying AI to the internal services of government

A professional public service is supported by intuitive and efficient internal services. Some of these services directly service Canada's democratic institutions, such as access to information or responses to the questions of parliamentarians. Others are in place to ensure that the public service itself is functioning smoothly, fostering a positive work environment and securing public assets.

6.3.1. Information Management

From white papers such as this one, to briefing notes, presentations, data sets, and other analysis, the GC is sitting on a vast trove of data, structured and unstructured, tagged and untagged. Traditional means of using this data has been limited to specific, machine-readable formats, but advances in semantic analysis have unlocked the potential for information in text format to be mined for insights as well. Now machine-usable information can be gleaned from text, audio, or video.

This technology can be used for a variety of applications, such as analyzing social media reaction to government policy or events; summarizing past briefings or approaches to maintain institutional memory; or automatically creating documentation trails for internal audit purposes.

The power behind these applications offers the promise of AI eventually providing virtual librarian services. With properly structured and tagged text data, a policy analyst will be able to more easily sort through and summarize past approaches to a problem, or find what is being done in other institutions. Having a smarter content management system understand what an analyst is looking for will help ensure that policy options are driven by data and that corporate memory is retained, leading to greater institutional wisdom.

6.3.2. Automated Content Generation

Over the past several years, products have entered the market allowing for content, be it text, audio, or visual content, to be generated automatically. Systems have been deployed in the private sector to automatically produce newspaper articles, blog content, or marketing copy. One notable example of this technology has been at the Associated Press newswire, which is estimated to be able to generate 2,000 news articles a second. After several months of training, configuration, and maintenance, the system is now able to post stories without any human intervention at all. The "AI journalist" is capable of doing this because a) there was a dataset large enough for the computer to extract best practices, and b) most of these reports contain only factual information, with limited nuance.

There are potential applications for the business of government. This technology can likely be adapted to a number of government documents that are produced on a regular basis in large quantities that are often factual and follow a certain formula or template. While certainly incapable of making normative considerations, this technology can be useful to summarize and compare. For example, it would be able to write Ministerial correspondence, background sections of briefing or meeting scenario notes, background of Question Period notes, etc. This would allow human public servants to focus on analysis, policy lenses, considerations, and strategies for next steps.

Responsible Artificial Intelligence in the Government of Canada

6.3.3. People Management

AI is transforming the discipline of human resources management, whether to gauge and optimize productivity, or to match individuals to suitable jobs. The ability to scan through the information of thousands of candidates using a more precise and insightful method than static keyword searches can potentially lead to more effective hiring decisions. Understanding the skills and credentials of effective and ineffective employees can provide insight as to the attributes of an ideal candidate. This can improve overall organizational effectiveness, but also help an individual find a job they may be ideal for but may lack traditional qualifications.

Another HR application of AI is performance assessment and management. These tools measure an employee's effectiveness against certain criteria, such as delivering on projects or replying to stakeholder inquiries. Using these tools, a manager is able to have a dashboard of the productivity of employees and the current status of their projects.

These tools can bring ethical risks and must be deployed with great care. For many of these systems to work properly, a continuous volume of data must be collected about a person's productivity. This is tantamount to ongoing surveillance of the employee, something that could cause harm to the employee's mental health.¹² Deep and persistent AI supervision of employees may contribute to the very anxiety that reduces their effectiveness at work, which in turn may hinder them from changing jobs. Furthermore, this system would have to reflect the changing context of a job, such as busy or quieter periods of work (i.e. in media relations), or jobs that produce work that is difficult to easily quantify (i.e. policy advice).

Additionally, identifying optimal productivity may fail in certain cultural contexts, as some employees may work differently. A veteran, indigenous person, or someone born abroad may choose to work different hours, or using different techniques, which while effective, may be difficult to measure. An AI trained only on employees of European descent may not effectively evaluate an employee that is not. The systems would be required to consider the diverse accommodations that may be required for employees with certain disabilities.

At the current state of technology, AI systems should be prohibited from making unsupervised decisions about HR. When AI is generating recommendations for management, it is very important that employees be made aware of them in advance if at all possible, and be provided with the opportunity to access the information collected about them.

6.3.4. Security and Access Management

AI can be applied to the way institutions provide, review or revoke IT system and building authorizations by establishing baseline normal behaviour of staff and learning when certain activities seem out of the ordinary. It can provide a better alignment of IT security with operations and reduce the number of ad-hoc requests for access to a system. This can reduce the workload of IT administrators, allowing them to focus on user needs that are exceptional.

¹² <http://onlinelibrary.wiley.com/doi/10.1111/ntwe.12039/abstract>

Responsible Artificial Intelligence in the Government of Canada

AI-powered cybersecurity and access control can further assist by allowing the detection of user needs at a granular level within a very short time, allowing users to have permissions better suited for what their job actually requires. AI can also be used to optimize permissions in business continuity planning.

Finally, there have been advances in machine learning cybersecurity applications that are designed to identify threats earlier, including internal threats where a sudden change of behaviour raises concern. While AI offers great promise in cybersecurity, it should be viewed as a single layer of protection, and not a substitute for existing systems and processes.

6.3.5. Financial Management

Whereas standard data analytics can provide significant value to institutions by helping them understand patterns in their accounting, advances in machine learning and natural language processing have led to a variety of applications for more intelligent financial management.

For example, contract intelligence applications help organizations automate contract review by scanning for mistakes and suggesting corrections. Machine learning systems are also available to help organizations continuously monitor for fraudulent or misappropriated expenditures by learning typical expenditure behaviour in flagging potential anomalies.

7. Policy, Ethical, and Legal Considerations of AI

With all of the potential use cases offering to improve policy and services, enthusiasm for AI in government has been high. Unfortunately, improper application of this technology can lead to negative outcomes for users, from frustrating service experiences to being mistakenly denied eligibility for benefits.

While the use of AI offers a lot of promise in improving the efficiency of government, it is important to approach its use with a strong ethical foundation. Machine ethics have been debated for years, and the Government of Canada should learn from these groundbreaking discussions to ensure that this transformative technology best serves the interest of everyone living in Canada.

As these agents grow to operate in increasingly sophisticated spaces, they act on behalf of the Crown, and should be subject to similar values, ethics, and laws as public servants and adherence to international human rights obligations. Institutions should incorporate these ethical principles in their application of AI:

- 1. People should always be governed – and perceive to be governed – by people;**
- 2. AI systems deployed on behalf of government should be trained to reflect the *Values and Ethics of the Public Sector* as well as Canadian and international human rights obligations; they should be used to reinforce these values where possible;**

Responsible Artificial Intelligence in the Government of Canada

- 3. Organizations are accountable for the actions of AI systems, and should build systems that are auditable;**
- 4. Understanding the need to protect privacy and national security, AI systems should be deployed in the most transparent manner possible;**
- 5. Organizations should ensure that reliable contingencies are in place for when AI systems fail, or to provide services to those unable to access these systems;**
- 6. AI systems should be developed in a diverse team that includes individuals capable of assessing the ethical and socioeconomic implications of the system;**
- 7. AI systems should be deployed in a manner that minimizes negative impact to employees where possible, and should, where feasible, be created alongside the employees that will work with them.**

The Government of Canada is committed to incorporating international norms and standards in ethical design when applying AI or any autonomous system. The first step to preventing negative outcomes is to understand what they are and how they occur.

There is no “average” Canadian; this country consists of a population diverse in background and circumstance. There will be users with unique challenges that will test the rigour and limitations of algorithms deployed by government. Institutions need to account for exceptions, minimizing cases that fall through the cracks, and providing recourse for the inevitable failures of the system.

7.1. Data, Bias, and Rights

Every field of data entered is an investment for the future. That data will be examined, validated, and manipulated individually and in aggregate possibly thousands of times in the cycle of their life. Traditionally, data entry was viewed as an input cost to be minimized by many federal institutions, but as the world moves more towards data-driven decisions, organizations are centering data governance in their core operations. This has unfortunately revealed a lack of consistent quality in data holdings.

Many AI applications are only as effective as the quality and quantity of their input data. The first step for an institution wishing to deploy an AI application is to ensure that the necessary training data is available, representative of the problem that needs to be solved, is readable by machine, and that the organization has the legal authority to collect and use this data. It also means adopting a culture of good data practices, and investing in the people and systems necessary to create, store, protect, and use data effectively.

7.1.1. Prevention of Data Bias

AI systems are not neutral; they will learn the biases of its programmers and the datasets used to train it. While unintentional, this bias can have ramifications that could range from embarrassing to serious. Even data that is incorrectly entered or labeled can have knock-on effects that affect real people in real ways.

Responsible Artificial Intelligence in the Government of Canada

This can particularly affect vulnerable populations, of whom data has been collected historically with varying quantity and quality.

The ability to distinguish, predict, and learn means that AI is able to operate in a more abstract and probabilistic fashion than earlier forms of computing. To do this, AI needs to be trained with datasets and oriented towards preferable outcomes. Both the training process and the selection of preferable outcomes carry with it the bias of the humans that collected and tagged the data, as well as the programmers that designed the algorithm. The collection of some data can be imperfect due to social or cultural stigma; for example, suicides and sexual assaults in Canada are both underreported.¹³¹⁴ Even the choice of which datasets to use and which to reject may entrench bias into the decision, and can lead to different outcomes.

Without enough training, an AI will have difficulty achieving its task, or will do so in a way that could lead to misinterpretations of data. Data collected in a certain socioeconomic context will echo in the decision-making of algorithms. The responsible policy manager needs to ensure that this important context is added to the analysis, and that they understand potential ways that AI can interpret input data incorrectly. Even controlling for certain variables won't necessarily protect from bias, as it can be derived from other, correlated variables; for example, excluding ethnicity from analysis won't necessarily protect from bias if the system can infer ethnicity from another variable such as a name.

The results of data bias can be highly problematic. As AI applications are more widely dispersed throughout society, a number of these unintentional but notable biases have been uncovered. For example, an algorithm used to predict crime in the United States has been shown to reinforce discriminatory policing because the crime data upon which it was trained was collected disproportionately in African-American neighbourhoods.¹⁵ According to a study by Carnegie Mellon University,¹⁶ women tend to be shown job ads for high-paying jobs less often than men as a result of search algorithms, likely due to the fact that women have been disproportionately missing from these positions in the past.

Machines can't learn contextual policy objectives such as social equity or environmental stewardship without being taught that these goals – while maybe not the explicit goal of the system – are necessary trajectories to be taken into consideration.

Algorithms themselves can affect the systems that they are trying to assess through a feedback loop. For example, a recidivism model that determines early release from prison, but being in prison longer increases the probability of recidivism, creating a feedback loop that increases incarceration time.

In applications where machine vision or audition may be applied to individuals, it's important that people are not excluded by virtue of ethnicity, accent, or disability. Some rare disabilities may not appear in

¹³ <https://www.statcan.gc.ca/pub/82-624-x/2012001/article/11696-eng.htm>

¹⁴ <https://www.statcan.gc.ca/pub/85-002-x/2017001/article/14842-eng.htm>

¹⁵ <https://mic.com/articles/156286/crime-prediction-tool-pred-pol-only-amplifies-racially-biased-policing-study-shows#.sGIb3QeCM>

¹⁶ 2015 study using 1,000 simulated persons, Link: <http://www.cmu.edu/news/stories/archives/2015/july/online-ads-research.html>

Responsible Artificial Intelligence in the Government of Canada

training data at all, which could lead to negative outcomes for individual. For example, a border camera scanning for predictors of risk may misinterpret a “tic” of an individual with Tourette syndrome as suspicious. These can manifest in a diverse fashion, and should not cause this person to undergo secondary inspection every time they pass through the border.

Social media presents an unprecedented opportunity to understand what some Canadians happen to be talking about on some subjects, but this approach brings significant risks. First, many Canadians do not use social media, so views cannot necessarily be taken as representative of the population. For example, in 2016, only 22% of Canadians used Twitter.¹⁷ Second, national government or private social media firms are capable of marshalling thousands of social media accounts that are artificial, expressing whatever views that they are paid to express. These botnet campaigns can inflate the prevalence of certain perspectives. Without strong countermeasures to detect deliberate attempts to distort public discourse with botnets, social media data should never be assumed to have been produced exclusively by humans.

7.1.2. Data for Insights and Privacy Rights

While many of the privacy risks brought by AI are not fundamentally new, the magnitude of data collected and the ability to manipulate this data beyond what a human is capable of brings a new dimension to these risks.

Algorithms capable of gathering insights from unstructured data mean that the nature of the information that we collect changes. For example, a name is no longer just a name, but a data point in a wider pattern that can reveal ethnic background. The government did not explicitly collect ethnicity, but an algorithm could extract that information from the person’s name nonetheless. One’s address can reveal correlations with income, health outcomes, or likelihood to encounter crime. With only a few variables known about an individual, it is possible to extrapolate an entire portrait about a person that is potentially very personal, and surely more than the person intended to disclose. This extrapolated information is not verified to be true; it is simply resulting from a series of statistical correlations that imply it may be true.

This can have unintended consequences. As an illustrative example, suppose that an institution wanted to predict the most successful outcomes possible from a grant program. It trains an algorithm based on a variety of historic information on the companies that typically apply, their officers, and the outcomes of those grants several years, themselves determined using public sources. Surprisingly, the findings are that women who have undergone a change of name are at a high risk of their business is failing.

Collecting – or rather, deriving – this new personal information from the individual is not necessarily unethical; the test lies in how this insight is used. In this case, the institution should compensate for the bias to ensure that these women have an equal opportunity to receive the grant.

To the extent that AI uses personal information it must comply with the *Privacy Act* or other departmental privacy codes. For the purposes of the *Privacy Act* and Privacy Impact Assessments (PIA), AI does not

¹⁷ <http://www.digitalnewsreport.org/survey/2017/canada-2017/>

Responsible Artificial Intelligence in the Government of Canada

represent a new program, but a new suite of tools. Existing program PIAs should be updated to reflect these new tools, as well as the new data that may be collected and used for the program.

For the purposes of Privacy Impact Assessments (PIA), AI does not represent a new program, but a new suite of tools. Existing program PIAs should be updated to reflect these new tools, as well as the new data that may be collected and used for the program.

7.2. Transparency and Accountability

7.2.1. Accounting for the Actions of AI: The “Black box” Problem

A cornerstone of responsible government is that Ministers are accountable for the affairs of their portfolio institutions, enshrined in legislation and custom. This accountability flows down from them to Deputy Heads and others within institutions through a variety of authorities from law to regulation to Treasury Board policy.

If a human makes a decision that is challenged, we rely on his or her explanation, the data that they were exposed to, and the outcome to figure out what happened. The limitations to this approach reflect the limitations of humans; memory may be unreliable, notes can be incomplete, and human biases can affect recall. Society relies on human judgement partially because it is the only option available to us.

Recall that advanced machine learning methods used today, such as neural networks, involve breaking down a problem into thousands – if not millions – of small decisions in order to reach an outcome. Some of these decisions are explicitly coded into the algorithm, but others are learned by the algorithm on its own based on input data. Much like the human brain, it is difficult to understand the entire decision-making process in detail made by the machine’s artificial neural network. This is known as the “black box” problem.

If an algorithm needs to be examined for whatever reason to determine the decision making process, hundreds of thousands of lines of code may need to be reviewed. Even then, it might be very difficult to reproduce exact results, or determine exactly why an outcome occurred. In the context of a neural network, examining each artificial neuron may not provide a sufficient understanding of the decision-making process.¹⁸

Invalid or biased decisions by algorithms tend to exist due to incomplete or biased datasets; therefore institutions should focus on the quality and completeness of training data, testing and audit findings of the applications, and any operating parameters. Systems should be continually tested and audited to ensure that the outputs still meet the original intention.

¹⁸ This is a rapidly-evolving area of research. Statistical and cryptographic techniques (e.g. Merkle trees) have been suggested to resolve this problem by creating an audit trail of decisions

Responsible Artificial Intelligence in the Government of Canada

Technology may be able to solve some of these problems; for example, there are tools in development¹⁹ that can trace and describe how a decision was made using a neural network.²⁰ TBS will have to continually monitor the development of this technology to ensure that transparency-enhancing techniques are adopted. Ultimately, how much explainability is required will largely be determined by jurisprudence. It is expected that this will be a higher test than what is expected from a human case manager. That said, institutions must ensure that decision-making algorithms provide enough detail around an explanation to understand why it was made for reasons of administrative and legal oversight like that carried out by the Information Commissioner, the Privacy Commissioner, the Auditor General, and the Courts.

If the government has to make decisions based on models that they don't understand or have access to, it hands some decision-making power to a private company with a black box. It is important that institutions have full understanding of the tools that they are using for decision support. To manage this risk, institutions may have to develop algorithms using internal or contracted resources, and maintain ownership of the input data and the algorithms used to make decisions. As well, institutions will have to retain all data used to train an AI for the duration that the AI is in use.

7.2.2. Social Acceptability

The enduring nature of Canadian democracy rests in the provision of good government to its citizens and residents. The Government of Canada is exploring the use of powerful tools at a time when trust in public institutions is low,²¹ and when a minority of Canadians feel that new technologies will do more good than harm.²² Not only do government AI applications need to be effective, but the population needs to perceive them as effective as well for them to be legitimate additions to, or substitutes for, human officials. If AI is going to make decisions, recommendations, or help design policy, there needs to be a sufficient level of social trust that these systems work, and work to the population's benefit. If the trust and support does not exist, then these tools will fail.

Trust will be built over time, assuming that the rules surrounding the use of these tools are transparent, and that appropriate information is available to users about how they work.

7.3. AI and the Law: An Emerging Landscape

AI will likely challenge current legal paradigms, although its actual impact on the law is still uncertain. In Canada, the technology is not comprehensively regulated, and there are few cases involving AI that have gone to court. If a government institution implements an AI solution that collects, uses, discloses or

¹⁹ <https://qz.com/1022156/mit-researchers-can-now-track-artificial-intelligences-decisions-back-to-single-neurons/>

²⁰ See for example: Pang Wei Koh, Percy Liang; Proceedings of the 34th International Conference on Machine Learning, PMLR 70:1885-1894, 2017 - Link: <http://proceedings.mlr.press/v70/koh17a.html>

²¹ <https://www.edelman.com/trust2017/trust-in-canada/>

²² Ipsos Canada Next, Public Perspectives. October 2017.
<https://www.ipsos.com/sites/default/files/ct/publication/documents/2017-10/public-perspectives-canadanext-2017-10-v1.pdf>

Responsible Artificial Intelligence in the Government of Canada

retains personal information various requirements of the federal *Privacy Act* may come into play as well as additional requirements found in applicable program or departmental statutes.

The use of AI in the government will undoubtedly have legal implications that range across many diverse areas of law, including Administrative Law, Privacy Law, Cyber-security Law, Intellectual Property Law, Crown Liability, Charter and Human Rights Law, Procurement Law, Employment Law, and the Law of Evidence.

Legal issues will be raised at each stage of use of AI, from its development to its deployment. For this reason, it will be very important to ensure that its use protects people's fundamental rights and that ethical and legal standards are considered at each stage of the use of AI, from the earliest stages of development onwards. To this effect, institutions should engage their institution's legal services unit as early as possible in the design and development of their project.

7.4. Technical Considerations

With the rapid expansion of AI-driven applications in the last five years, there have been moves to ensure that technical standards are established for interoperability, best practices, and increasingly, safety as well. International standard-setting organizations such as the IEEE Standards Association, OpenAI, and ISO have established technical standard working groups to bring to the tools, methods, and practices associated with algorithm development and implementation. With an aim to adopting international best practices, TBS will be closely following the developments of these standards with an aim to provide guidance to institutions soon as appropriate.

Analytics and AI projects are inherently multidisciplinary and therefore will cross many teams and branches in an organization. Federal institutions exploring this technology should be sure to include their IT and data management teams from the design stage onward. These applications will need to fit consistently in the enterprise architecture to allow for secure connectivity with client relationship management software and other data repositories both within your organization and beyond. Chatbots, for example, may need to interact with your web presence if they include navigational capability. While the communications or policy function within an institution may be the business owners of this technology, it's important to remember that there may need to be connections with an institutions client relationship management software, or need to draw data from other organizations.

Institutions will need access to the right tools in a timely fashion. Data science teams should have access to the required software and servers without undue delay. Access to secure cloud computing services rated to protected B will also be required.

7.4.1. Cybersecurity considerations

Many of the cyber security implications to AI are similar to those of other critical systems and government. However, there are some new threats to take into consideration.

Responsible Artificial Intelligence in the Government of Canada

Databases containing training data must be properly secured from intrusion. Even if the intruder cannot extract meaning from the data, changing figures in training data can lead to changing the outputs of the algorithms. The effects of these changes could be wide-ranging, unanticipated, and difficult to detect.

As mentioned above, deep learning chatbots can be taught to provide incorrect or inappropriate information. Keyword blacklists only go so far to prevent inappropriate behavior, because a chatbot doesn't necessarily need to use explicitly bad language to provide bad service. The training data provided to chat box need to be properly secured, so that an intruder cannot, for example, reroute a link from a government service to of malicious clone used to steal personal information, or be tricked into behaving in inappropriate ways.

An intruder may want to shut down an automated decision support system that an institution relies upon. Like many critical systems, redundancies will need to be put in place to prevent lengthy outages to critical systems. This should include that at least some human staff are retained and properly trained for manual backups.

8. Rethinking a Post-AI Enterprise

AI tools will be transformative for government, but only if institutions are ready to deploy them. The first step is investing in data science and business intelligence capacity within your Institution. This includes skilled personnel, tools, storage, and data governance mechanisms. These investments should be led by a skilled Chief Data Officer, who has control and access to data sets throughout the organization, an understanding of related data holdings of others, and a direct reporting relationship to the Chief Information Officer of the organization or a suitable Assistant Deputy Minister.

The second step is to ensure that the deployment of AI applications come with a multidisciplinary and diverse project team. Having a mix of social scientists, ethicists, data scientists, change management professionals and user experience designers from a variety of backgrounds is a potent defense against data biases and other risks preventing your organization from reaching success. Easy and accessible user experience is vital to the uptake of these tools, regardless if they are not directly used by the public.

8.1. New Approaches to the Workforce

Advances in artificial intelligence have brought with them a very public discussion about the future of work and the role that knowledge workers will play in the economy. To date, there are examples of task automation in the private sector that lead to significant staff reductions; conversely, there are examples where no workforce reductions were required. AI has shown itself to be a valuable suite of tools that can either exponentially increase productivity, or eliminate routine tasks, allowing for humans to perform more valuable work for the organization. That said, this transition does not happen on its own, and anxiety that can result from perceived imminent automation can have a real impact on employee

Responsible Artificial Intelligence in the Government of Canada

wellbeing and productivity.²³ As institutions look to automate work, it is important that they choose approaches that maximize utility to the organization while at the same time minimizing the potential for staff reductions.

At the same time, post-automation and AI government will require new skills from existing staff and new types of staff from the labour market. Federal institutions will need to attract data scientists, invest in up-to-date tools that they can use, and access to relevant data sources.

Institutions will need to ensure that their policy analysis and development teams understand how to access, interpret, and manipulate data relevant to their work, and to have access to the skills development resources that they need to grow within their field. New employees to a team should be provided with context on how their data has been collected and used in the past, and how stakeholders typically view this collection and use.

Federal institutions should also be reminded that some collective bargaining agreements contain specific sections on workforce adjustment due to technological change. To ensure that these requirements are honored, TBS recommends that unions and non-represented staff alike are engaged early in the planning phases. Staff and unions will be useful partners to help automate processes in a way that is both most useful to the user as well as least affecting of positions.

8.2. Evolving How Government Works

Departmental internal audit, central agencies, and agents of Parliament together play a role to ensure that programs are designed in a manner that is compliant with policy and best practices. This robust system of oversight ensures that institutions are accountable to ministers, Parliament, and the public. In an era of increasingly data-driven policy recommendations and autonomous systems, does the government have the right tools to oversee its business?

There is no organization that currently exists with the clear mandate and capability to respond to complaints of data biases or algorithmic design. Canadians and parliamentarians alike will need an obvious contact point to manage these issues, and oversight organizations provided the tools necessary to do this job. While algorithms cannot be completely transparent to avoid fraud, an oversight body staffed with the required expertise could be provided access when required.

TBS will require to do further research on models of governance that could provide the necessary oversight and guidance to Federal institutions. This can range from an *ad hoc* federal “Automation Advisory Board” comprising of internal and external experts to a more formal and permanent body with staff. Regardless of the model chosen, the body would have the ability to review automated decision-making by any methodology, and provide advice to ministers during the design - but especially prior to Cabinet approval of projects - on ethical design of AI-driven programs and services.

²³ The Economist, “Automation and Anxiety,” June 2016. Link: <http://www.economist.com/news/special-report/21700758-will-smarter-machines-cause-mass-unemployment-automation-and-anxiety>

Responsible Artificial Intelligence in the Government of Canada

9. Conclusion

AI applications have emerged as useful tools for institutions to include in their policy, program, and service development process. They can bring exponential power to government, but must be applied where it makes sense. While AI applications can be applied to many projects, but should only be considered if there is a reasonable value proposition from its use. Simply adding a machine learning component to a project will neither guarantee its success nor be the sole guarantor of smarter policymaking. Introducing AI into a program brings risks that need managing, and requires staff capable of managing it.

AI's complex and multidisciplinary nature demands that federal institutions work together, sharing talent and best practices, to leverage knowledge and avoid duplication. It means working with other orders of government, and with research institutions and non-governmental organizations within Canada and beyond.

Government is also fundamentally about people and relationships. Machines cannot substitute for empathy, and even the best analytics will find outliers that cannot be forgotten. TBS actively encourages institutions to explore this technology for the benefit of the populations that we serve. Ethical and responsible design of these systems will drive a virtuous cycle of acceptance, which in turn will drive further development.

As a next step, TBS will begin to examine its policy suite to ensure that existing guidance is useful to institutions that will implement AI applications in their organizations. Where appropriate, standalone guidance will be considered as well. As a rapidly evolving technology, TBS will require ongoing engagement to ensure that policy is reflective of technological capabilities so that institutions can continually make best use of what AI has to offer.

Responsible Artificial Intelligence in the Government of Canada

10. Acknowledgements

This paper was drafted using an “open” approach, with contributors welcome to engage in its development from conception to completion. Treasury Board of Canada Secretariat would like to thank the numerous and ongoing contributions from participants in all sectors.

Responsible Artificial Intelligence in the Government of Canada

[8] <http://onlinelibrary.wiley.com/doi/10.1111/ntwe.12039/abstract>

[16] <http://www.ekospolitics.com/index.php/2014/01/looking-back-and-looking-forward-part-2/>

Will contributors be thanked in this version?

Protected – Solicitor-client Privilege

EVERGREEN: last updated August 9, 2018

Contacts: Nathalie Sarault and Julia Betts
TBS Legal Services

Service Strategy
Project Overview for TBS Legal Services
(Legal Project Management)

1. Project Description
<ul style="list-style-type: none">The objective of the Strategy is to improve quality, timeliness, and efficiency of federal government service delivery to Canadians, with three desired results: 1) Client-driven Design and Delivery; 2) Easy Online Services; and 3) Seamless Delivery.
2. Governance
<p>Client:</p> <ul style="list-style-type: none">TBS role (CIOB): develop broad policy objectives, oversee the implementation of the Strategy and establish performance standards. TBS main clients: Sonya (Acting Senior Director) and Simon Pomel (Director).Other departments (ESDC (Service Canada), VAC, CRA, CBSA, AAFC, PSPC, SSC, IRCC, CFIA, ISED): implement specific projects. <p>DOJ:</p> <ul style="list-style-type: none">Virtual team of legal counsel (listed below) from relevant departments/sections where members work together as needed and keep each other informed on a regular basis, including through standing meetings, to address legal services requirements from the clients.TBS Legal Services to coordinate consultation within the virtual team of legal counsel, as needed.
3. Key Milestones Achieved
<ul style="list-style-type: none">
4. Timelines and Next Steps

s.69(1)(g) re (a)

1

Protected – Solicitor-client Privilege

EVERGREEN: last updated August 9, 2018

Contacts: Nathalie Sarault and Julia Betts
TBS Legal Services

<ul style="list-style-type: none"> • June 26, 2018: ADM SFI to present Work Plan • Mid-July 2018 – TBS to send departments a “Questions Paper” to inform next steps for the 3 “quick wins”, in preparation for a workshop • Early August – workshop with departments to discuss quick wins • August – tasking to departments to identify use cases (i.e. services requiring improvement) • September – December: in-depth review of use cases and plan for resolution (i.e. improved service state)
<p>5. Key Legal Considerations</p>
<ul style="list-style-type: none"> • Identification of legal issues and solutions for implementing a new Service Strategy, mainly with respect to privacy, mandate and authorities.
<p>6. DOJ Stakeholders</p>
<ul style="list-style-type: none"> • TBS LSU (GOPM): Nathalie Sarault, Julia Betts, Lisa Bambrick • CIPL: Magali Clervoix for legal advice on Service Strategy issues (note: Daniel Caron is the DOJ contact for the Privacy Act review and for any policy linkages to be made between the Privacy Act review and the Service Strategy) • AAFC LSU/CFIA LSU: Andrea Horton • CBSA LSU: Mélanie Laframboise • CRA LSU: Janice Palmer • ESD (Service Canada) LSU: Nathalie Déziel • VAC (part of ESDC LSU): Susan Carruthers • IRCC LSU: Michelle Mann • ISED LSU: Beatrice Van Rутten • PSPC LSU (SSC): Kimberley Lewis • Stats Can : Isabelle Stears • Commercial Law Group : Gabriella Signorini • (Placeholder for LSB, CAILS (for now, consult with Marie Lasnier, as discussed with Julie Wellington), Finance, PSC, other DOJ units)
<p>7. Key DOJ Documents (completed or under preparation)</p>
<ul style="list-style-type: none"> • [REDACTED] • Service Strategy – Legal Landscape (ongoing)

s.23

Protected – Solicitor-client Privilege

EVERGREEN: last updated August 9, 2018

Contacts: Nathalie Sarault and Julia Betts
TBS Legal Services

- Legislative inventory (ongoing)
- [REDACTED] s.69(1)(g) re (f)
- SIN Paper (draft)
- ISA Paper (draft)
- Keys opinions (see G:\LEGAL - JURIDIQUE\ADVISORY WORK - CONSEILS\SERVICE STRATEGY - 8773000\DLSU opinions on Service Strategy)

8. Related Files

- [REDACTED]
- OneGC – vision of the GoC operating as one; it includes DXP (digital exchange platform, formerly the interoperability platform) and “tell us once” – the concept of Canadians being able to provide information only once to the GoC rather than having to repeat it when dealing with different departments.
- Digital Policy – includes enabling digital services to Canadians and the work on artificial intelligence.
- Data Strategy – to better manage GoC information as a valuable asset; it includes data relevant for providing services to Canadians, for statistical purposes, and for informing policy and legislative/regulatory decisions.
- [REDACTED]
- DESDA – ESDC

9. Other notes of interest

s.23

**Pages 106 to / à 118
are withheld pursuant to section
sont retenues en vertu de l'article**

69(1)(g) re (e)

**of the Access to Information Act
de la Loi sur l'accès à l'information**

NLS BOARD OF DIRECTORS MEETING

June 26, 2018 – 3:00 p.m. EDT
Room 542, 50 O'Connor Street

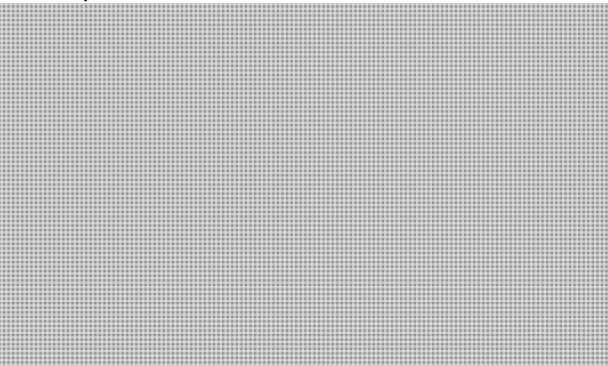
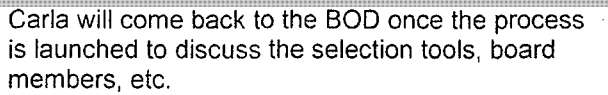
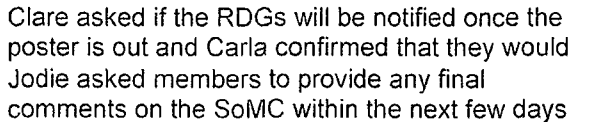
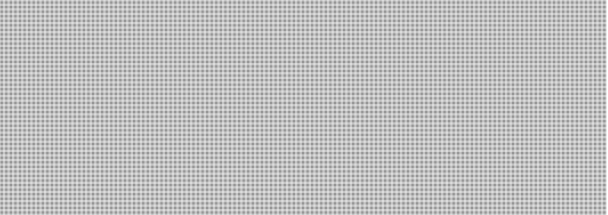
Record of Decision

3:00 – 3:05 pm / 15h00 à 15h05	1.	<p>Welcome / Bienvenue</p> <p>Review of Record of Decision / Révision du compte rendu de décision</p> <ul style="list-style-type: none"> Geoff asked that any requested changes to the ROD from the June 19, 2018 meeting be sent to Jenna Melrose 	Geoffrey Bickert
3:05 – 3:35 pm / 15h05 à 15h35	2.	<p>Open Government / Le gouvernement ouvert</p> <ul style="list-style-type: none"> 	<p>s.21(1)(a) s.21(1)(b)</p> <p>Kristine Allen / Nora Ballantyne / Karen Wallace</p>

s.21(1)(a)

s.21(1)(b)

		<ul style="list-style-type: none"> • [REDACTED] • [REDACTED] • [REDACTED] 	
3:35 – 3:55 pm / 15h35 à 15h55	3.	<p>Staffing Dashboards / Tableau de bord en dotation</p> <ul style="list-style-type: none"> • Melanie presented the Staffing Dashboards for the NLS and the Department as a whole • Some highlights of the presentation: <ul style="list-style-type: none"> ○ The reports will be run twice per year to track trends ○ Regional breakdowns are also available upon request ○ [REDACTED] ○ Of the current workforce, 80% are indeterminate and 20% are term ○ [REDACTED] ○ [REDACTED] ○ [REDACTED] 	Melanie Stethem
3:55 – 4:05 pm / 15h55 à 16h05	4.	LP-05 / LP-05	Carla Lyon

		<ul style="list-style-type: none"> • Carla presented the SoMC which incorporated the feedback she's received from the BOD, as well as HR • She noted that the closing date will be amended before it's posted – it will remain open for approximately 1 month after posting (rather than 2 weeks) •  •  •  • Carla will come back to the BOD once the process is launched to discuss the selection tools, board members, etc. • Clare asked if the RDGs will be notified once the poster is out and Carla confirmed that they would • Jodie asked members to provide any final comments on the SoMC within the next few days so that we can launch the process 	<p>s.21(1)(a) s.21(1)(b)</p>
4:05 – 4:10 pm / 16h05 à 16h10	5.	<p>Round Table / Tour de table</p> <ul style="list-style-type: none"> • James noted a lack of practitioners serving as members on the Working Group to the Task Force on Artificial Intelligence. He felt that the NLS should have 2 representatives on the Working Group. He will send an email to the BOD asking for volunteers; names are to be submitted by July 5th. Yves noted that the regions should suggest someone from the Tax Section who is currently involved in the Tax Foresight pilot AI project •  • Clare asked if an item could be put on an upcoming agenda providing an update on the National collective process for ECs. Jodie confirmed this will be placed on a future agenda • Marie-Claude confirmed she has now replaced Anik Picard 	<p>All / Tous</p>

Page 122

**is withheld pursuant to sections
est retenue en vertu des articles**

21(1)(a), 21(1)(b)

**of the Access to Information Act
de la Loi sur l'accès à l'information**

RÉUNION DU CONSEIL D'ADMINISTRATION DU SNC

Le 26 juin 2018 – 15 h (HAE)
50, rue O'Connor, salle 542

Compte rendu des décisions

15 h à 15 h 05 / 3:00 – 3:05 pm	1.	<p>Bienvenue / Welcome</p> <p>Révision du compte rendu des décisions / Review of Record of Decision</p> <ul style="list-style-type: none"> Geoff demande que tous les changements sollicités pour le compte rendu des décisions de la réunion du 19 juin 2018 soient envoyés à Jenna Melrose. 	Geoffrey Bickert
15 h 05 à 15 h 35 / 3:05 – 3:35 pm	2.	<p>Le gouvernement ouvert / Open Government</p> <ul style="list-style-type: none"> 	<p>s.21(1)(a) s.21(1)(b)</p> <p>Kristine Allen / Nora Ballantyne / Karen Wallace</p>

s.21(1)(a)

s.21(1)(b)

		<ul style="list-style-type: none"> • • • 	
15 h 35 à 15 h 55 / 3:35 – 3:55 pm	3.	<p>Tableau de bord en dotation / Staffing Dashboards</p> <ul style="list-style-type: none"> • Melanie présente le tableau de bord en dotation pour le SNC et le Ministère dans son ensemble. • Voici les points saillants de sa présentation : <ul style="list-style-type: none"> ○ Les rapports seront présentés deux fois par année pour suivre les tendances. ○ Les ventilations par région sont également disponibles sur demande. ○ ○ De la main-d'œuvre actuelle, 80 % des employés sont nommés pour une période indéterminée et 20 % pour une période déterminée. ○ ○ 	Melanie Stethem

		<ul style="list-style-type: none"> ○ [REDACTED] 	<p>s.21(1)(a) s.21(1)(b)</p>
<p>15 h 55 à 16 h 05 / 3:55 – 4:05 pm</p>	<p>4.</p>	<p>LP-05 / LP-05</p> <ul style="list-style-type: none"> • Carla présente l'ECM avec l'intégration des commentaires qu'elle a reçu du CA et des RH. • Elle souligne que la date butoir sera modifiée avant que l'ECM soit affiché – il restera ouvert pendant environ un mois après l'affichage (au lieu de deux semaines). • [REDACTED] • [REDACTED] • [REDACTED] • Carla reviendra au CA lorsque le processus sera amorcé pour discuter des outils de sélection, des membres du conseil, etc. • Clare demande si les DGR seront avertis lorsque l'annonce sera publiée et Carla confirme qu'ils le seront. • Jodie demande aux membres de présenter leurs commentaires finaux sur l'ECM dans les prochains jours pour que l'on puisse démarrer le processus. 	<p>Carla Lyon</p>
<p>16 h 05 à 16 h 10 / 4:05 – 4:10 pm</p>	<p>5.</p>	<p>Tour de table / Round Table</p> <ul style="list-style-type: none"> • James souligne le manque de professionnels œuvrant comme membres au sein du groupe de travail de la commission d'étude sur l'intelligence artificielle. Selon lui, le SNC devrait avoir deux représentants au sein du groupe de travail. Il enverra un courriel au CA pour demander des volontaires; les noms devront être soumis avant le 5 juillet. Yves souligne que les régions devraient suggérer quelqu'un de la Section du droit fiscal qui participe actuellement au projet pilote sur l'IA de Tax Foresight. • [REDACTED] 	<p>Tous / All</p>

s.21(1)(a)

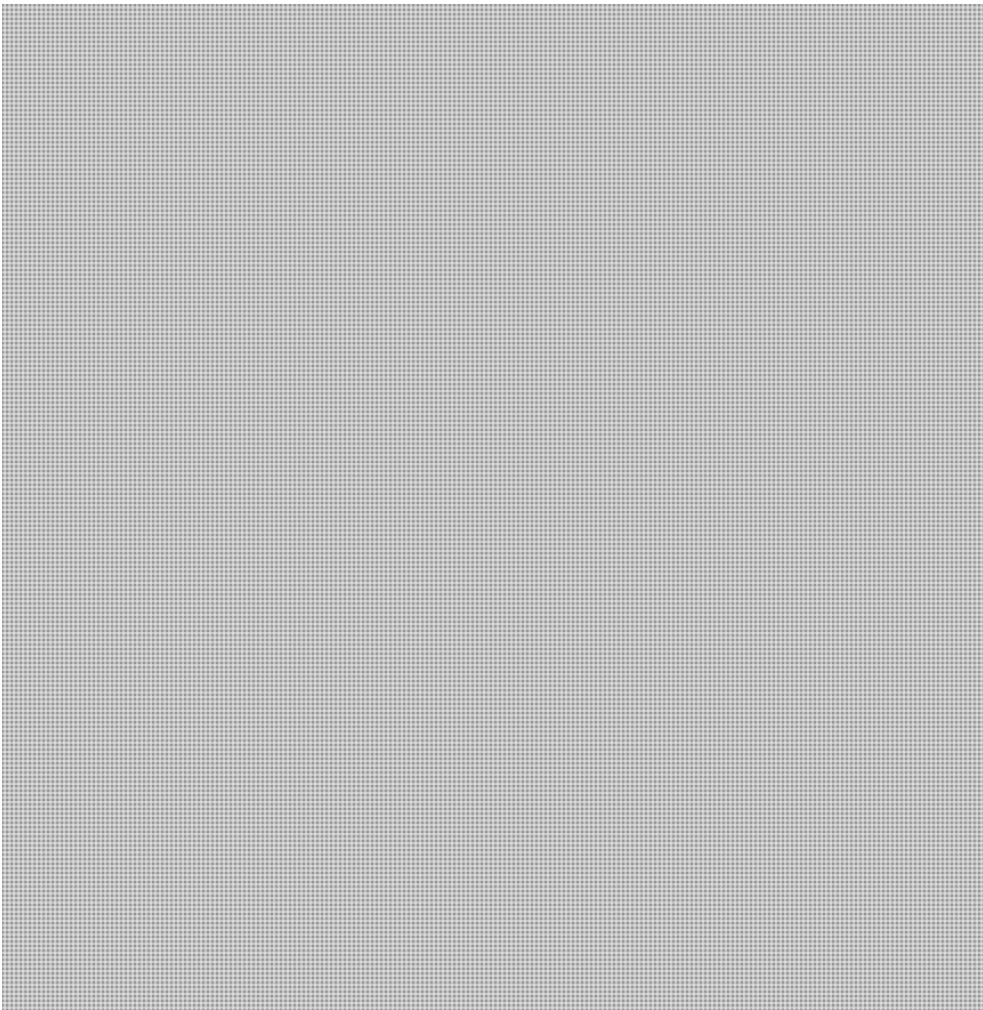
s.21(1)(b)

			<ul style="list-style-type: none">• Clare demande si un point peut être ajouté à un prochain ordre du jour pour offrir une mise à jour au sujet du processus collectif national pour les EC. Jodie confirme que ce point sera ajouté à un futur ordre du jour.• Marie-Claude confirme qu'elle a maintenant remplacé Anik Picard.•	

Last update: August 2, 2018

Key contacts: Michelle Mann (IRCC LSU) and Nathalie Sarault (TBS LSU)

**THE WORKING GROUP TO THE TASK FORCE ON ARTIFICIAL
INTELLIGENCE
DRAFT 2018 08 02
Terms of Reference**



s.21(1)(a)

s.21(1)(b)

s.23

**Pages 128 to / à 130
are withheld pursuant to sections
sont retenues en vertu des articles**

21(1)(a), 21(1)(b), 23

**of the Access to Information Act
de la Loi sur l'accès à l'information**

Thompson, Chris

From: Mann, Michelle P. (CIC)
Sent: 2018-Jul-11 12:51 PM
To: Stringham, James; Tardif, Richard; Sarault, Nathalie; Sarault, Nathalie (TBS)
Subject: RE: AI Working Group draft questions in relation to the AI WG Terms of Reference

Not too late - thanks very much!

Michelle Mann

Senior Counsel, Legal Services (IRCC)
Immigration, Refugees and Citizenship Canada / Government of Canada
Michelle.Mann@cic.gc.ca / Tel: 613-437-6366

Avocat-conseil, Services juridiques (IRCC)
Immigration, Réfugiés et Citoyenneté Canada / Gouvernement du Canada
Michelle.Mann@cic.gc.ca / Tél.: 613-437-6366

From: Stringham, James [mailto:James.Stringham@justice.gc.ca]
Sent: July 11, 2018 12:22 PM
To: Mann, Michelle <Michelle.Mann@cic.gc.ca>; Tardif, Richard <Richard.Tardif@justice.gc.ca>; Sarault, Nathalie <Nathalie.Sarault@justice.gc.ca>; Sarault, Nathalie (TBS) <Nathalie.Sarault@tbs-sct.gc.ca>
Subject: RE: AI Working Group draft questions in relation to the AI WG Terms of Reference

You are welcome Michelle.

s.21(1)(a)

s.21(1)(b)

I neglected to provide input on the rest of the ToRs, so I hope it is not too late for me to do so now.

s.23

On page 3, the draft reads:

I consulted with our ADAGO, with the thought that the ADAGO does a lot of coordination and tracking of litigation to ensure the ADAG and DM are kept informed of developments, and that it might be redundant for the Working Group to set up tools to do the same. With that in mind, the ADAGO suggested adding "Where appropriate" to the "c" clause as well. I am sure the NLS reps on the Working Group will let the WG know when the ADAGO already is tracking something...

Best regards,

James

From: Mann, Michelle [mailto:Michelle.Mann@cic.gc.ca]
Sent: July-11-18 10:40 AM
To: Stringham, James <James.Stringham@justice.gc.ca>; Tardif, Richard <Richard.Tardif@justice.gc.ca>; Sarault, Nathalie

<Nathalie.Sarault@justice.gc.ca>; Sarault, Nathalie (TBS) <Nathalie.Sarault@tbs-sct.gc.ca>
Subject: RE: AI Working Group draft questions in relation to the AI WG Terms of Reference

Thanks for your very helpful input James!

Michelle

Michelle Mann

Senior Counsel, Legal Services (IRCC)
Immigration, Refugees and Citizenship Canada / Government of Canada
Michelle.Mann@ic.gc.ca / Tel: 613-437-6366

s.19(1)
s.21(1)(a)
s.21(1)(b)
s.23

Avocat-conseil, Services juridiques (IRCC)
Immigration, Réfugiés et Citoyenneté Canada / Gouvernement du Canada
Michelle.Mann@ic.gc.ca / Tél.: 613-437-6366

From: Stringham, James [<mailto:James.Stringham@justice.gc.ca>]

Sent: July 6, 2018 1:52 PM

To: Tardif, Richard <Richard.Tardif@justice.gc.ca>; Mann, Michelle <Michelle.Mann@ic.gc.ca>; Sarault, Nathalie <Nathalie.Sarault@justice.gc.ca>; Sarault, Nathalie (TBS) <Nathalie.Sarault@tbs-sct.gc.ca>

Subject: RE: AI Working Group draft questions in relation to the AI WG Terms of Reference

From: Stringham, James

Sent: July-05-18 2:28 PM

To: Tardif, Richard <Richard.Tardif@justice.gc.ca>; Mann, Michelle P. (CIC) <Michelle.Mann@ic.gc.ca>; Sarault, Nathalie <Nathalie.Sarault@justice.gc.ca>

Subject: RE: AI Working Group draft questions in relation to the AI WG Terms of Reference

Richard, Michelle and Nathalie,

Best regards,

James Stringham
Directeur et Avocat général | Director and General Counsel

Preuve électronique et soutien aux litiges | eDiscovery and Litigation Support
Secteur national du contentieux | National Litigation Sector
Ministère de la Justice Canada | Department of Justice Canada
360 Albert Street, Tower 1, 14th floor
Ottawa (Ontario) K1A 0H8
Téléphone | Telephone (613) 948-3477/ Télécopieur | Facsimile (613) 941-6822
BlackBerry: (613) 277-0559

james.stringham@justice.gc.ca

s.19(1)
s.21(1)(a)
s.21(1)(b)
s.23

CONFIDENTIALITY NOTICE: The contents of this electronic mail message are confidential and strictly reserved for the sole use of its intended recipients. This message may contain information protected by solicitor-client privilege. If you have received this message in error, please notify the sender immediately and delete the original message as well as all copies. Any disclosure, copying, distribution or reliance on the contents of the information is strictly prohibited. Thank you.

AVIS DE CONFIDENTIALITÉ: Le contenu de ce courriel est confidentiel et strictement réservé à l'usage des personnes auxquelles il s'adresse. Ce message peut contenir de l'information protégée par le secret professionnel. Si vous avez reçu ce message par erreur, contactez immédiatement l'expéditeur et détruisez le message original et toute copie. Veuillez noter qu'il est strictement interdit de divulguer, reproduire ou diffuser le contenu de ce message ou de prendre des mesures en conséquence. Tout nouvel envoi, reproduction ou usage de ce message par une personne autre que le destinataire est strictement interdit. Merci.

From: Tanyan, Caroline On Behalf Of Tardif, Richard

Sent: June-22-18 6:59 PM

To: McCombs, Mark (HRSDC) <mark.mccombs@hrsdc-rhdcc.gc.ca>; Zimmerman, Corinne <Corinne.Zimmerman@justice.gc.ca>; Wojcik, Nicolas <Nicolas.Wojcik@justice.gc.ca>; Stringham, James <James.Stringham@justice.gc.ca>; Wellington, Julie <Julie.Wellington@justice.gc.ca>; Fobes, Caroline (CIC) <Caroline.Fobes@cic.gc.ca>; Mann, Michelle P. (CIC) <Michelle.Mann@cic.gc.ca>; Wong, Robert (FIN) <robert.wong@canada.ca>; McDonald, Susan <Susan.McDonald@justice.gc.ca>; Reaney, Jennifer <Jennifer.Reaney@justice.gc.ca>; Signorini, Gabriella <Gabriella.Signorini@justice.gc.ca>; Zadro, Matthew <Matthew.Zadro@justice.gc.ca>; Dubrulle, Louis-Philippe <Louis-Philippe.Dubrulle@justice.gc.ca>; Thompson, Chris <Chris.Thompson@justice.gc.ca>; Topshee, Dugald <Dugald.Topshee@justice.gc.ca>; Sarault, Nathalie (TBS) <Nathalie.Sarault@tbs-sct.gc.ca>

Subject: AI Working Group draft questions in relation to the AI WG Terms of Reference

Dear AI Task Force members,

I would like to thank again Michelle Mann and Nathalie Sarault for preparing a draft of the Terms of Reference for the DOJ AI WG (latest draft is attached).

In order to finalize the TOR, they would benefit from our comments on the following points:



I was wondering if you could provide Michelle and Nathalie with your comments on the TOR and the 5 questions prior to next Thursday. Thank you also for cc'ing me on your replies.

Regards.

Richard

s.21(1)(a)

s.21(1)(b)

s.23

M^e Richard L. Tardif
Director General and Senior General Counsel |
Directeur général et avocat général principal
Legal Practices Sector | Secteur des pratiques juridiques
Department of Justice Canada | Ministère de la Justice Canada
275 Sparks Street, 10th floor, SAT-10012 | 275, rue Sparks, 10^e étage, TSA-10012
Ottawa (Ontario) K1A 0H8
Tel. | Tél. : 613-952-3816
richard.tardif@justice.gc.ca

CONFIDENTIALITY NOTICE: The contents of this email are confidential and reserved for the sole use of the intended recipients. Any use, disclosure or copying of the information is strictly prohibited. This email may also contain information protected by solicitor-client privilege. If you receive this message in error, please notify the sender immediately and destroy the original message and all copies. Thank you.

AVIS: Ce courriel contient des renseignements confidentiels dont l'usage est réservé exclusivement à la personne à laquelle il est destiné. Toute utilisation, divulgation ou reproduction de cette information est strictement interdite. Ce courriel peut aussi contenir de l'information privilégiée qui est protégée par le secret professionnel de l'avocat-client. Si vous recevez cette communication par erreur, veuillez en aviser l'auteur immédiatement et détruire l'original et toute copie. Merci.